Mechanisms of DNA Methylation Defects at the *IGF2/H19* Imprinting Centre in Patients with Foetal Growth Disorders

A thesis submitted in fulfilment of the requirements for the degree of Master of Applied Science (Medical laboratory Science)

Mansur Ennuri Shmela

B.V.Sc.

School of Medical Sciences

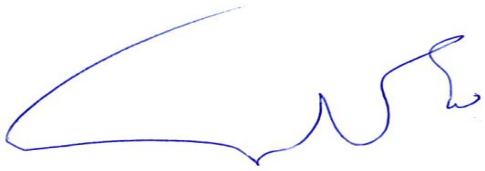
College of Science, Engineering and Health

RMIT

May 2009

DECLARATION

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.



Mansur Ennuri Shmela

09/05/2009

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisors:

- **Dr. Christine Gicquel,** for her continued assistance and great guidance throughout the entire period of my study and without her support, this thesis would not exist.

- **Dr. Peter Roche,** for his immediate help and enormous support when I needed.

- **Ass/Prof. Assam El-osta,** who gave me this opportunity to work in his laboratory and for his great encouragement.

- Heartfelt thanks to all members of the epigenetic laboratory for their excellent technical assistance and kind support.

- Special thanks to Dr. Julie Demars for her grateful assistance and support during the last year.

- Many thanks to other people at Baker institute for their technical assistance, especially Dr. Tye Dawood.

- Special thanks to my friends here in Australia for their kind support.

- My thanks are due to my country (Libya) for providing with a scholarship to undertake my study.

- Finally, I am grateful and thankful to my family in Libya for their love and kind words of encouragement, specially my parents, who never stop praying for me to be successful in my academic career and my life.

IN THE NAME OF ALLAH THE MOST GRACIOUS THE MOST MERCIFUL

To my father and mother,

Ennuri and Fatima,

To my brothers and sisters,

To my sister's son and daughter,

Waleed and Sajeda,

I dedicate this simple work

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ALS	Acid-labile subunit
bp	Base pairs
BWS	Beckwith-Wiedemann syndrome
Cdkn1c	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)
CpG	Cytosine-phosphate-Guanine
CTCF	CCCTC-binding factor
DMD	Differentially methylated domain
DMR	Differentially methylated region
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides
EDTA	Ethylene diamine tetraacetic acid solution
H19	Imprinted maternally expressed transcript (non-protein coding) gene
H3K27	Histone 3 lysine 27
H3K4	Histone 4 lysine 4
H3K9	Histone 3 lysine 9
HAT	Histone acetyl transferase
HMT	Histone methyl transferase
hrs	Hours
ICR	Imprinting control regions
IGF	Insulin-like growth factor
IGF1	Insulin-like growth factor 1
IGF1/ Igf1	Insulin-like growth factor 1 gene
IGF1R	IGF1 receptor
IGF2	Insulin-like growth factor 2
IGF2/ Igf2	Insulin-like growth factor 2 gene
IGF2R	IGF2 receptor
IGF2r/Igf2r	*IGF2* receptor gene
IGFBP	Insulin-like growth factor binding protein
InsR	Insulin receptor

Kb	Kilo base
KCNQ1	Potassium voltage-gated channel, KQT-like subfamily, member 1 gene
M	Molar
MBD	Methyl-CpG binding domain
MeCP2	Methyl-CpG binding protein 2
mg	Milligram
MgCl2	Magnesium chloride
min	Minute
ml	Milliliter
mM	Millimolar
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
ncRNA	Non-coding RNA
ng	Nanogram
P	Promoter
PCR	Polymerase chain reaction
PRC	Polycomb repressive complex
rpm	Round per minute
SDS	Sodium dodecyl sulfate solution
sec	Second
SNP	Single-nucleotide polymorphism
SRS	Silver-Russell syndrome
TE	Tris EDTA buffer
TRIS	Trishydroxymethylaminomethane
UPD	Uniparental disomy
V	Volt
X-gal	5-bromo-4-chloro-3-indolyl beta-galactoside
ZF	zinc finger
°C	Degree Celsius
µg	Microgram
µl	Microlitre
µM	Micromolar

ABSTRACT

The imprinted expression of the *IGF2* and *H19* genes is controlled by the imprinting control region 1 (ICR1) located at chromosome 11p15.5. This methylation-sensitive chromatin insulator works by binding the zinc-finger protein CTCF in a parent-specific manner. CTCF binds the unmethylated maternal allele and is required for preventing *de novo* methylation at ICR1.

DNA methylation defects involving the ICR1 *IGF2/H19* domain result in two growth disorders with opposite phenotypes: an overgrowth disorder, the Beckwith-Wiedemann syndrome (ICR1 gain of methylation in 10% of BWS cases) and a growth retardation disorder, the Silver-Russell syndrome (ICR1 loss of methylation in 60% of SRS cases). Little information is available regarding the mechanism of ICR1 DNA methylation defects. Several deletions removing part of ICR1 (1.4 to 2.2 kb) have been described in a few familial BWS cases with dominant maternal transmission.

In order to evaluate precisely the incidence of ICR1 mutations, we investigated, by long range PCR and sequencing, 21 BWS patients (including two brothers) with ICR1 gain of methylation and 16 SRS patients with ICR1 loss of methylation.

No mutation of the seven CTCF binding sites was detected in the familial BWS cases. Two additional cases of constitutional genetic lesions were identified in BWS patients with apparently-sporadic forms. One patient was identified with a 8 bp deletion within the B3 repeat, 116 bases 3' of the CTCF binding site 4. Another patient was identified with a 1.8 kb deletion which eliminates CTCF binding sites 2 and 3. A single-nucleotide variation was identified in a SRS patient.

Our data showed that ICR1 deletions, including new small deletions, account for apparently sporadic forms of BWS with ICR1 gain of methylation. ICR1 deletions are associated with a high incidence of Wilms' tumour, making their molecular diagnosis particularly important for genetic counseling and tumor surveillance.

1 INTRODUCTION

1 INTRODUCTION

1.1 Genomic Imprinting

The majority of genes in diploid organisms are expressed from both alleles. However, a number of genes are expressed from one of the two parental alleles; these genes are known as imprinted genes. In the same cell, some imprinted genes are expressed from paternally inherited alleles (maternally imprinted) and others are expressed from maternally inherited alleles (paternally imprinted) (SOLTER 2006).

In the early 1980s, two different pronucleus transplantation studies demonstrated that the paternal and maternal chromosomes do not have equivalent functions and that both genomes are essentially required for normal mammalian development (MCGRATH and SOLTER 1984; SURANI *et al.* 1984). One decade later, the maternally expressed *Igf2r* gene (Insulin-like growth factor-2 receptor) was the first gene found to be expressed in a parent of origin-dependent manner and defined as an imprinted gene (BARLOW *et al.* 1991). The same year, two other imprinted genes, the *Igf2* and *H19* genes on the mouse chromosome 7 were shown to display reciprocal imprinting and to be paternally and maternally expressed respectively (BARTOLOMEI *et al.* 1991; DECHIARA *et al.* 1991). Many other imprinted genes have since been discovered and more than 80 imprinted genes have been reported in mammals so far (http://www.geneimprint.com).

Genomic imprinting is highly conserved among mice, ruminants and humans (DELAVAL and FEIL 2004). A conflicting theory of genomic imprinting suggests that imprinting arose because of a genomic tug-of-war between the two parents where maternally expressed genes inhibit growth while paternally expressed genes promote growth (WEIDMAN 2007) . Imprinted genes often code for proteins involved in foetal growth and development and placental function (GICQUEL *et al.* 2008). Other imprinted genes are also expressed in the brain and influence

the developmental processes of the brain as well as brain function and behaviour (WILKINSON *et al.* 2007). Therefore, any disturbance of genomic imprinting results in many growth and mental disorders and learning disability (GICQUEL *et al.* 2008; WILKINSON *et al.* 2007).

1.1.1 Characteristics of imprinted genes

Imprinted genes display common features. Almost all known imprinted genes are **organized in clusters** which usually include both imprinted and nonimprinted genes (WEIDMAN 2007). Clusters of imprinted genes are regulated by single major *cis*-acting elements which are known as imprinting control regions (ICRs). These regions are CpG (Cytosine-phosphate-Guanine) rich regions up to several kilobases in length and are differentially methylated between the two parental alleles **(differentially methylated regions (DMRs))** (LEWIS and REIK 2006)**.** ICRs are also marked by histone modifications (phosphorylation, methylation, acetylation and ubiquitinylation) (KOUZARIDES 2007). As a result, ICRs are essential for the regulation of the expression of imprinted genes in an allele-specific manner.

ICRs regulate imprinted gene expression by two main mechanisms; **chromatin insulator** and **non-coding RNA gene (ncRNA)** (SHA 2008; WEIDMAN 2007) (i) Chromatin insulator mediates imprinted gene expression through special DNA sequences at ICRs. These sequences bind specific proteins, such as the CTCF protein (see 1.2.4.2), which prevents the interaction between gene promoters and enhancers and therefore suppresses gene expression (LEWIS and REIK 2006). (ii) Another mechanism of regulation of imprinted gene expression involves ncRNA genes. Each imprinted cluster expresses at least one large ncRNA that displays reciprocal imprinted expression relative to the neighbouring coding genes. Some of these ncRNA genes are transcribed in an antisense orientation relative to the protein-coding genes. These long ncRNA genes act as *cis*-acting silencers. Imprinted gene clusters also contain small regulatory ncRNAs, and particularly

microRNAs which act as *trans*-acting regulators (ROYO and CAVAILLE 2008).

Another feature associated with imprinted genes is a **difference in the timing of allele replication** between the two parental alleles. The replication timing pattern has been examined for some imprinted genes in the mouse and it has been shown that the paternal allele replicates earlier than the maternal allele (SHA 2008).

1.1.2 Epigenetic marks of imprinted genes

Homologous genes, which carry the same genetic information, can be expressed differentially according to the parental origin. In imprinted genes, the maternal and paternal alleles convey different gene expression as a result of different epigenetic marks without affecting the genetic code (TRASLER 2006). Therefore, the epigenetic marks tell the cell which parental copy of imprinted genes must be activated. DNA methylation and histone modifications are the two common epigenetic marks which are involved in imprinted gene regulation.

1.1.2.1 DNA methylation

DNA methylation is one aspect of epigenetic modifications in which the 5' position of cytosine in the CpG dinucleotide acquires a methyl group (SHA 2008). This epigenetic mark is inherited throughout replication and is essential for the mammalian development (LI 2002; OKANO *et al.* 1999). Regarding imprinted genes, the DNA methylation pattern undergoes a cycle which is discussed in section (1.2.3).

Addition and maintenance of a methyl group at the 5' position of cytosine within CpG dinucleotides are mediated by DNA methyltransferases (Dnmts) (table I). Dnmts which introduce a methyl group at previously unmethylated CpG sites are known as ***de novo* methyltransferases** (Dnmt3 family) (TRASLER 2006). Dnmt3a and Dnmt3b are essential for *de novo* methylation in early postimplantation embryos but they have no effect on maintenance of imprinted

methylation patterns (BESTOR 2000; HATA *et al.* 2002). Dnmt3l is a protein sharing homology with Dnmt3 a and b to carry out *de novo* methylation of maternally imprinted genes in oocytes (BOURC'HIS *et al.* 2001; HATA *et al.* 2002).

The other Dnmts group is involved in the maintenance of the pre-existing methylation, this group is known as **maintenance methyltransferases** (Dnmt1 group). Dnmt1 copies the pre-existing methylation pattern during DNA replication into the new DNA strand (HOWELL *et al.* 2001; LI *et al.* 1992). Dnmt1o is an oocyte isoform of Dnmt1 and is crucial for the maintenance of DNA methylation at imprinted loci in the fourth S phase of embryogenesis (HOWELL *et al.* 2001). Following the implantation, the oocyte specific Dnmt1o disappears and Dnmt1 protein maintains the methylation during development (HOWELL *et al.* 2001; LI *et al.* 1992; TRASLER *et al.* 1996).

Normal DNA methylation status is required for controlling differential expression of the paternal and maternal alleles of imprinted genes. Aberrant DNA methylation is linked to a number of imprinting-related disorders in humans (GICQUEL *et al.* 2008; GICQUEL *et al.* 2005; TRASLER 2006). The mechanistic role of DNA methylation in association with gene expression remains largely unclear. In imprinted regions, spreading of DNA methylation along ICRs is suggested to coordinate gene expression of a cluster (SHA 2008; TURKER 1999). Recently, long-range chromatin interaction or “looping”, which is modulated by both DNA methylation and histone modifications, has been proposed for coordination of imprinted gene expression (LOPES *et al.* 2003; MURRELL *et al.* 2004).

Table I: Role of the different DNA methyltransferases

DNA methyltransferase	Function	Specific role in genomic imprinting	References
Dnmt1	Maintenance of the pre-existing methyl pattern during DNA replication.	Dnmt1o isoform is required for maintenance of methylation at paternally and maternally imprinted loci during the fourth embryonic S phase.	(LI *et al.* 1992) (HOWELL *et al.* 2001)
Dnmt3a	*De novo* methylation during the embryonic developmental process. Redundancy with Dnmt3b.	Methylation of maternal ICRs in association with Dnmt3l.	(BESTOR 2000) (HATA *et al.* 2002)
Dnmt3b	*De novo* methylation during the embryonic developmental process. Redundancy with Dnmt3a. Specifically required for methylation of pericentromeric satellite repeats.		(BESTOR 2000)
Dnmt3l	Absence of proper DNA methyltransferase activity.	Methylation of maternal ICRs in association with Dnmt3a.	(BOURC'HIS *et al.* 2001)

1.1.2.2 *Histone modifications and chromatin structure*

The other epigenetic marks involved in regulation of genomic imprinting are histone modifications. DMRs at imprinted loci are not only marked by DNA methylation but also by chromatin modifications that participate to a different chromatin organization between the parental alleles (SHA 2008) (Fig. 1). DNA wraps around a histone protein octamer which consists of two copies of each histone, H2A, H2B, H3 and H4 to form the nucleosome. Each histone contains a functional domain (N-terminal tail) that protrudes from the nucleosome and acts as a target site for enzymes and proteins (WORKMAN and KINGSTON 1998). N-terminal tails of histones can be modified by different post-translational modification mechanisms such as methylation, acetylation, phosphorylation and

ubiquitinylation. These mechanisms are catalysed by different enzymes (KOUZARIDES 2007). Histone modifications have different effects on the chromatin structure (BERGER 2007). Histone methylation is associated with active or inactive chromatin depending on the site of methylation. For example, addition of a methyl group, by histone methyl transferases (HMTs), to the histone 3 lysine 4 (H3K4) is associated with uncondensed (active) chromatin. On the other hand, H3K9 and H3K27 methylation are associated with condensed (inactive) chromatin. Lysine residues can be mono-, di- and trimethylated and provide further functional diversity to each methylated lysine residue (BERGER 2007; KOUZARIDES 2007; STRAHL and ALLIS 2000). Acetylation marks seem to be less complicated than methylation marks and commonly link to an active state at a given region due to a relaxed chromatin structure and increased access of transcriptional factors. For instance, H3 and H4 acetylation, mediated by histone acetyl transferases (HATs), are associated with transcriptional activation (LI *et al.* 2007).

Although histone modifications are found independent of DNA methylation, there is clear evidence that links DNA methylation and histone modification to control gene expression (LEWIS *et al.* 2004). In different studies, DNA methylation was found to control histone modifications and vice versa. For example, DNA methylation recruits methylated CpG binding complexes that contain enzymes able to add a repressive methylation mark (H3K9) to histones (SARRAF and STANCHEVA 2004).

The histone code of the murine ICR1 has been extensively characterized by Han *et al*. Rather than, displaying active histone modifications on the unmethylated allele and repressive histone modifications on the methylated allele, the allelic distribution of the histone marks vary along the ICR1 domain. Active marks (H3K4 dimethylation, H3K4 trimethylation, H3K9 acetylation) are associated with the maternal *H19* gene and the paternal *IGF2* gene and the allelic

distribution is the opposite for repressive marks (H3K9 dimethylation, H3K27 trimethylation). Regarding the imprinting center itself, repressive marks are equally distributed on both parental alleles. Interestingly, deletion abolishing CTCF binding sites within ICR1 on the maternal allele results in the "paternalization of the maternal allele"(HAN *et al.* 2008).

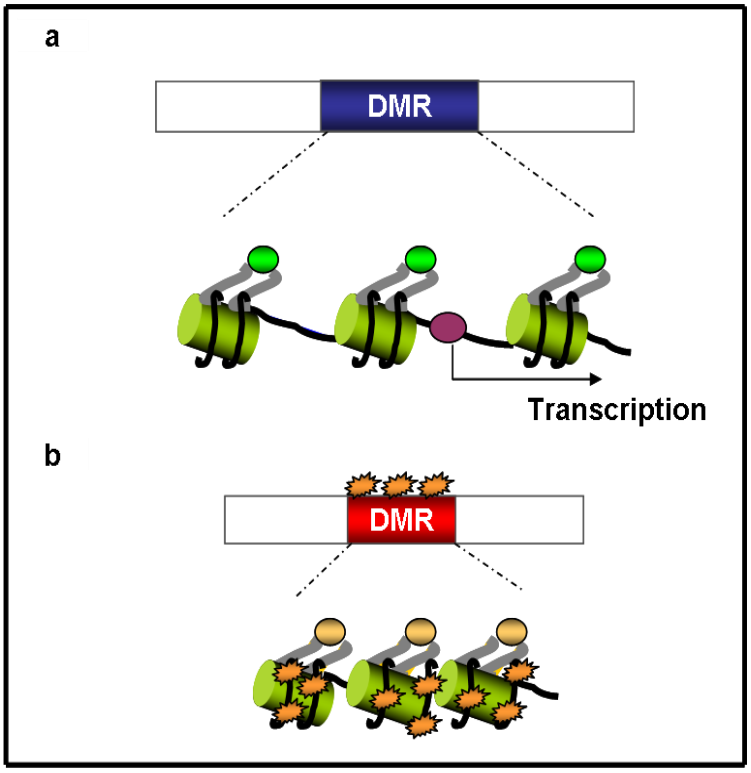


Figure 1: Epigenetic marks and chromatin structure at DMRs. **a**) Active epigenetic marks on an active allele. Unmethylated DMR and active histone modifications (green circles) open the chromatin and allow the access of transcriptional factors to start transcription. **b**) Inactive epigenetic marks on an inactive allele. Methylated DNA (orange stars) and inactive histone modifications (beige circles) condense the chromatin and deny the access of transcriptional factors to start transcription (adapted from (OZANNE and CONSTANCIA 2007)).

1.1.3 The imprinting cycle

The epigenetic marks of imprinted genes change in characteristic ways during the life cycle of the organism. Three different stages are involved in the imprinting cycle: **erasure** of imprints, **establishment** of new

parent-specific imprints and **maintenance** of imprints throughout development (Fig. 2 and 3) (SOLTER 2006; TRASLER 2006).

Erasure occurs in primordial germ cells in the gonads before their differentiation. All pre-existing imprints are erased during genome-wide demethylation in germ cells (HAJKOVA *et al.* 2002; LI 2002; SHA 2008; TADA *et al.* 1998). For example, the methylation of ICR1 at the *IGF2*/*H19*–ICR1 locus is erased in foetal spermatogonias and reappears later during spermatogonial differentiation in the adult testis (KERJEAN *et al.* 2000). Erasure of imprints is thought to be active but the enzymes involved in this developmental process are still unknown (TRASLER 2006).

After erasure of previous imprints, *de novo* methylation begins in the male and female germ lines and some imprints are **established** in oocytes and others in sperm. At this stage, ICRs are targeted by two epigenetic marks, DNA methylation and histone modifications (DELAVAL *et al.* 2007). Dnmt3 enzymes are key enzymes for the acquisition of gamete-specific DNA methylation in both gametocytes (see 1.1.2.1) (TRASLER 2006). However, it is still unclear why some ICRs are methylated in the paternal germline and other in the maternal germline. Some proteins such as CTCF (which is not expressed in the testis) prevents the acquisition of methylation at *H19* differentially methylated domain (DMD) in oocytes (ENGEL *et al.* 2006; FEDORIW *et al.* 2004; PANT *et al.* 2003).

Following fertilization, the two parental genomes are exposed to a wave of demethylation followed by a wave of *de novo* methylation, both of which are resisted by imprinted loci (MORGAN *et al.* 2005). The mechanisms involved in the **protection** of ICRs at this stage are not elucidated (TRASLER 2006). The Dnmt1 and its isoform Dnmt1o are essential for the maintenance of methylation at imprinted loci (HOWELL *et al.* 2001; LI *et al.* 1992; TRASLER *et al.* 1996). Methyl-CpG binding domain (MBD) proteins as well as histone modifications might also play a role in the maintenance of ICRs DNA methylation by affecting

chromatin conformation (DELAVAL and FEIL 2004; MORGAN *et al.* 2005). At the *H19 DMD* in mice *(IGF2*/*H19*–ICR1 locus in humans), CTCF is required for the protection of the maternal allele from gain of methylation (ENGEL *et al.* 2006; PANT *et al.* 2003).

The timing pattern of DNA replication of imprinted genes is also found to be erased in the germ line and then a parent-specific replication timing is established in late stage of gametogenesis and maintained throughout development (SHA 2008; SIMON *et al.* 1999).

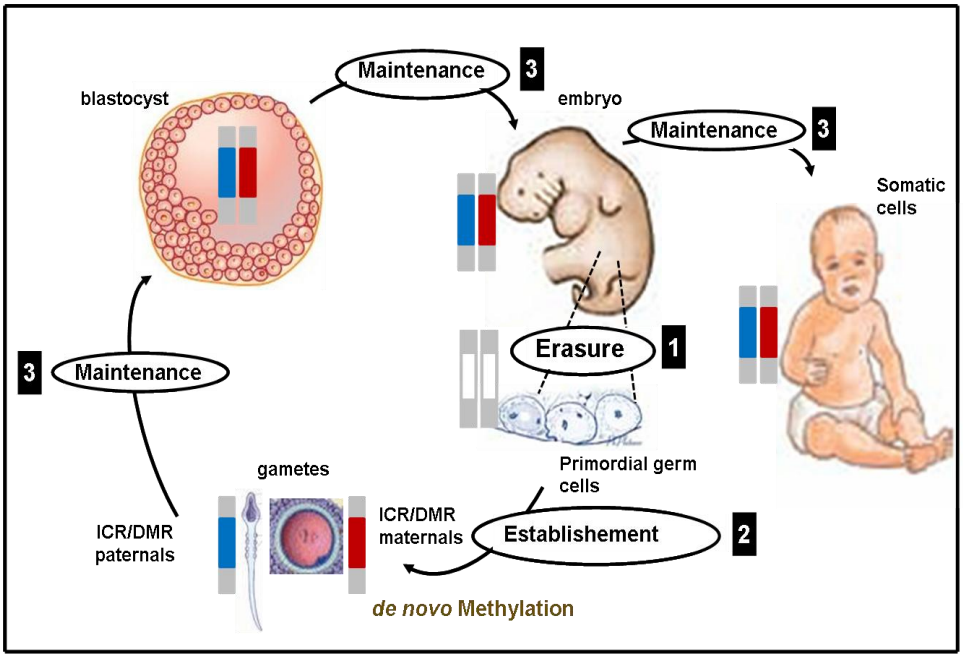


Figure 2: The imprinting cycle. In primordial germ cells, the previous imprinting marks are erased (1) at a very early stage and they will be re-established (2) at a later stage during gametogenesis according to individual sex (maternal-type marks in the oocytes and paternal-type marks in sperm). During early development, the imprinting marks should be maintained (3) while there is a wave of global demethylation followed by a wave of a *de novo* methylation. The imprinting marks are then maintained in somatic cells during development (Adapted from (REIK and WALTER 2001)).

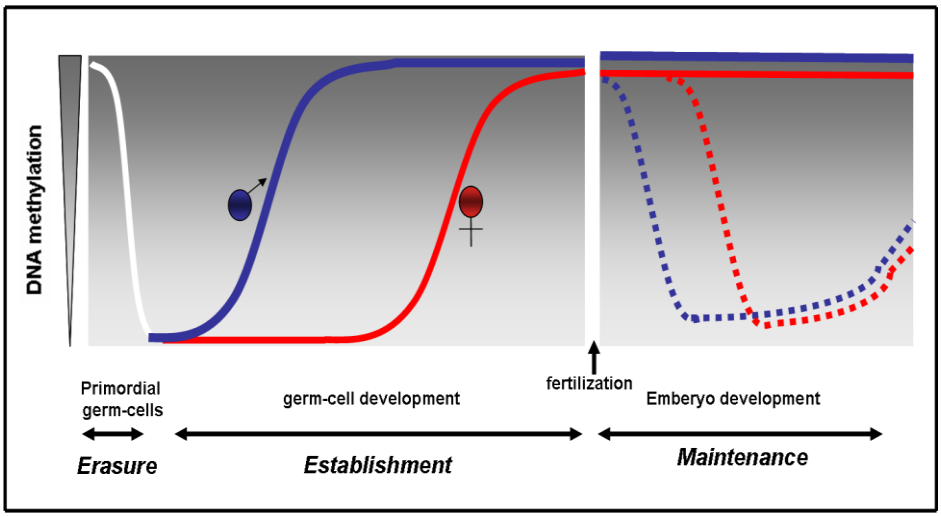


Figure 3: DNA methylation dynamics during germ cell and preimplantation development: erasure in primordial germ cells (white), establishment of paternal (blue) and maternal (red) imprinting marks during spermatogenesis and oogenesis and maintenance of imprints after fertilization; the dotted lines represent the wave of demethylation of the paternal and maternal genomes which imprinted genes are protected from (Adapted from (REIK and WALTER 2001)).

1.2 The *IGF2/H19*-ICR1 domain

Human chromosome 11p15.5 contains a cluster of imprinted genes that play an important role in foetal and placental growth. This region comprises two neighbouring imprinted domains, the *IGF2/H19* and the *KCNQ1* domains, each of them under the control of its own imprinting center, ICR1 and ICR2 respectively (Fig. 4) (DELAVAL *et al.* 2006; GICQUEL and LE BOUC 2006; REIK and WALTER 2001). The *IGF2/H19*-ICR1 domain includes two imprinted genes with reciprocal imprinting, *IGF2* and *H19,* which are paternally and maternally expressed, respectively (IDERAABDULLAH *et al.* 2008). The mouse distal chromosome 7 region is orthologous to the human 11p15.5 region and displays the same imprinted genes (DELAVAL *et al.* 2006).

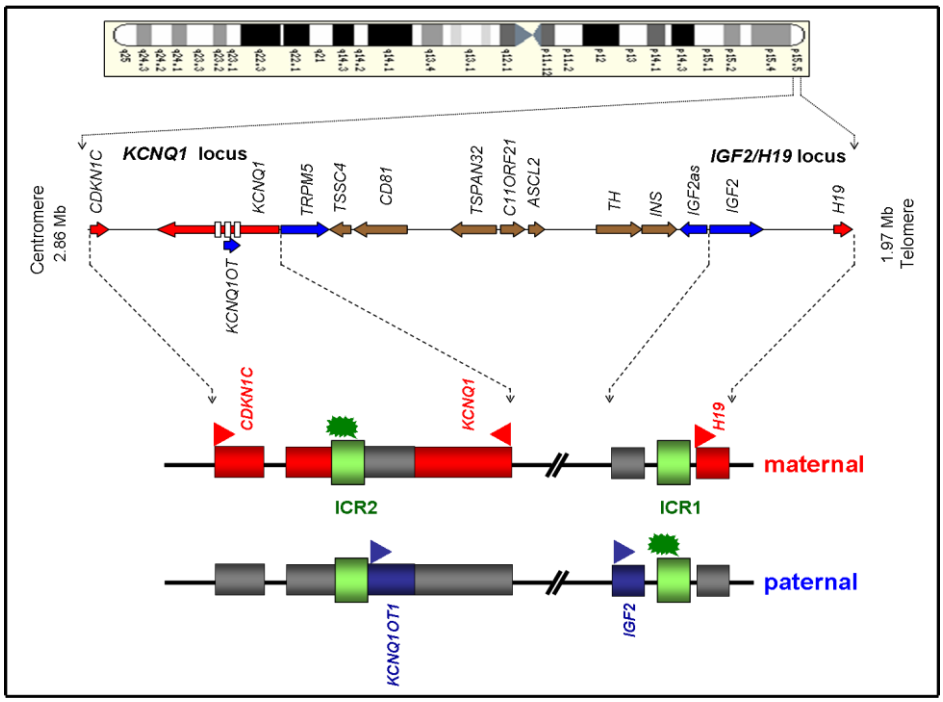


Figure 4: The imprinted cluster on human chromosome 11p15.5. It includes two imprinted domains, *IGF2/H19*-ICR1and *KCNQ1*-ICR2. Genes which are expressed from both alleles (not imprinted) are represented in brown, genes expression from the maternal allele are represented in red and genes expression from the paternal allele are represented in blue colour. ICR1 is methylated only on the paternal allele (green stars) and ICR2 is methylated only on the maternal allele.

1.2.1 Overview of the IGF system

The insulin-like growth factor system (IGF) is one of the most important endocrine and paracrine systems that regulates the complex process of foetal and placental growth. The IGF system consists of two ligands (IGF1 and IGF2), two IGF receptors (IGF1R and IGF2R) and a family of 6 binding proteins (IGFBPs) from 1 to 6 (Fig. 5) (GICQUEL and LE BOUC 2006).

The biological effects of IGF1 and IGF2 are conveyed through the IGF1R receptor. IGF2 can also bind the insulin receptor (InsR). IGF2R controls IGF2 levels negatively by targeting IGF2 for lysosomal degradation and therefore inhibits foetal growth. The majority of IGFs (~80%) bind to the IGFBPs and acid-labile subunit (ALS) to form a ternary complex in the blood stream. More than 20% of IGFs are

associated with one of the IGFBPs in a binary complex and less than 1% of the IGFs circulate in the free form (GICQUEL and LE BOUC 2006).

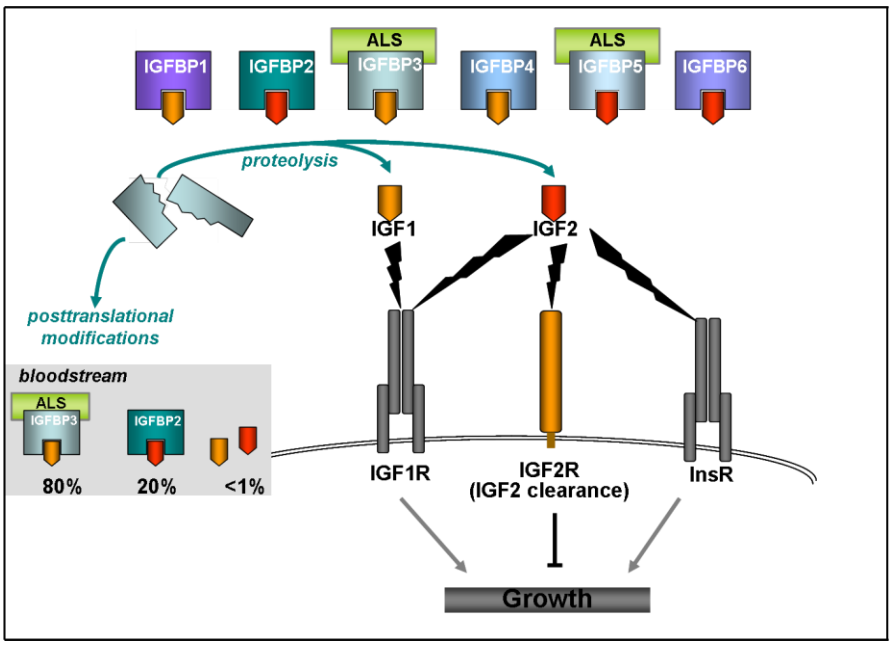


Figure 5: Schematic illustration of the insulin-like growth factor (IGF) system. The IGF system comprises IGF1, IGF2, IGF binding proteins (IGFBPs) 1 to 6, and the IGF receptors (IGF1R and IGF2R). IGF1 and IGF2 promote growth through interaction with IFG1R. IGF2 also promote growth through the insulin receptor (InsR). IGF2R is responsible for IGF2 clearance. The six IGFBPs bind both IGF1 and IGF2, but with different affinities. InsR = insulin receptor; ALS = Acid-labile subunit (Adapted from (GICQUEL and LE BOUC 2006)).

Several studies have been conducted to demonstrate the effect of IGFs on foetal and placental growth in mice (Fig. 6) (BAKER *et al.* 1993; GICQUEL and LE BOUC 2006). Inactivation of *Igf1* or *Igf2* results in foetal growth retardation (60% of normal body weight) and more severe growth retardation is caused by deletion of *Igf1r* (45% of normal body weight). Double knockouts of *Igf2* and *Igf1* or *Igf2* and *Igf1r* result in a more severe growth retardation phenotype (30% of normal body weight). In contrast, *Igf2* transactivation results in foetal overgrowth (130% of normal body weight). Interestingly, knockout of the *Igf2* gene causes a reduction in placental size but neither *Igf1* nor *Igf1r* mutations affect placental size. All these results show that both IGFs are involved in foetal growth and that IGF2 also contributes to placental growth.

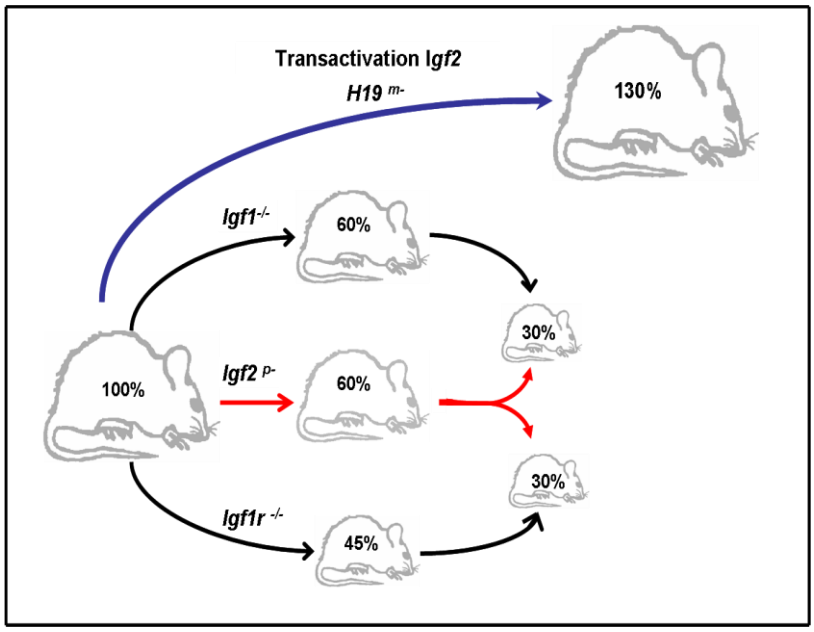


Figure 6: Effects of disruption of one or a combination of genes of the IGF system on foetal growth in mice, expressed as a percentage of normal body weight. m– = Maternally disrupted allele; p– = paternally disrupted allele; –/– = both alleles disrupted (Adapted from (GICQUEL and LE BOUC 2006)).

1.2.2 The IGF2 gene

As previously discussed, the *IGF2* gene plays a crucial role in foetal and placental growth. *IGF2* is imprinted in most tissues and is expressed from the paternal allele. The human *IGF2* gene consists of 9 exons and its expression is controlled by 5 different promoters (P0 to P4) (Fig. 7) (FOWDEN *et al.* 2006). During foetal life, the expression of *IGF2* is monoallelic in most tissues and driven by promoters P2, P3 and P4. After birth, the expression of the *IGF2* gene is biallelic in the liver and driven by promoter P1. In the remaining tissues, the expression remains monoallelic under the control of promoters P3 and P4 (FOWDEN *et al.* 2006). In addition, *IGF2* itself contains two DMRs (DMR0 and DMR2), which carry the same DNA methylation marks as the ICR1 imprinting center and are methylated on the paternal allele (MURRELL *et al.* 2008).

In the mouse, *Igf2* is also imprinted (paternally expressed) but there are differences between the human and mouse *IGF2* genes (Fig. 7). Firstly,

there are 4 promoters (P0 to P3) and no liver specific promoter. Secondly, the P0 promoter is specific of the placenta and is extremely important for placental transport (CONSTANCIA *et al.* 2002; FOWDEN *et al.* 2006). This promoter P0 is not specific of the placenta in humans. Moreover, the mouse *Igf2* gene displays three DMRs (DMR 0 to 2) and the mouse DMR1 behaves like the human DMR0 (MURRELL *et al.* 2008).

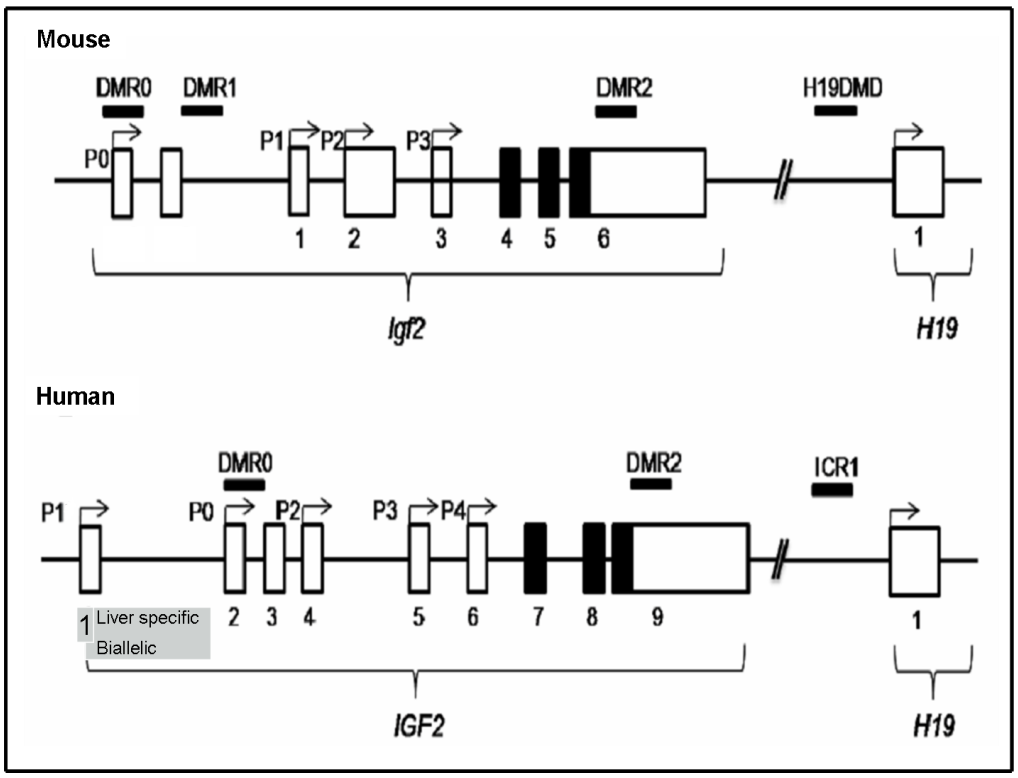


Figure 7: Schematic scale diagram of the mouse *Igf2/H19* and human *IGF2/H19*. P: promoter; DMR: Differentially Methylated Region; DMD: Differentially Methylated Domain; ICR: Imprinting Control Region. The coding region of *IGF2* is represented by black rectangles and the arrows represent start of transcription (Adapted from (MONK *et al.* 2006)).

1.2.3 The H19 gene

Clusters of imprinted genes usually consist of at least one ncRNA and one or more protein coding genes. The *H19* gene encodes a non translated RNA (BRANNAN *et al.* 1990) expressed only from the maternal allele in endodermic and mesodermic tissues (ZHANG and TYCKO 1992). The *IGF2* and *H19* genes have a reciprocal expression

(OHLSSON *et al.* 1994), under the control of the imprinting centre ICR1 (DELAVAL *et al.* 2006; GICQUEL and LE BOUC 2006; REIK and WALTER 2001). The expression of *H19 gene* is also regulated by a promoter which is differentially methylated, like ICR1, on parental alleles (GABORY *et al.* 2006). To understand the function of the *H19* gene, different deletions of this gene and the imprinting centre have been performed in mice. Mice in all these studies were viable and fertile, suggesting that the role of *H19* during development is not crucial. However, these mice with deletions exhibited overgrowth (GABORY *et al.* 2006).

Recently, the *H19* nonccoding RNA (ncRNA) has been shown to function as a primary micro RNA transcript that might act as a posttranscriptional down regulator of specific messenger RNAs (CAI and CULLEN 2007). More recently, a new transcript named *91H* RNA, which is expressed from the maternal allele within the *H19* gene region, has been found to maintain *IGF2* expression in *trans* (on the paternal allele), rather than affecting *H19* expression (BERTEAUX *et al.* 2008).

1.2.4 The epigenetic regulation of the reciprocal imprinting of the IGF2 and H19 genes.

1.2.4.1 *Structure and chromatin organization of the ICR1domain*

The human *IGF2*/*H19*-ICR1 domain (*H19* differentially methylated Domain (DMD) in mice) includes two reciprocally imprinted genes: the *IGF2* and the *H19* genes (Fig. 8). The reciprocal imprinting of the paternally expressed *IGF2* (OHLSSON *et al.* 1994; OHLSSON *et al.* 1993) and the maternally expressed *H19* (ZHANG and TYCKO 1992) genes depends on the differentially methylated region (ICR1) upstream from the *H19* gene (FREVEL *et al.* 1999; IDERAABDULLAH *et al.* 2008). This region is approximately 5 kilo-base (kb) in humans and 2 kb in mice and functions as an insulator (FILIPPOVA 2008; IDERAABDULLAH *et al.* 2008). In humans, the wild-type ICR1 domain is arranged in two repeat units and each of them has A repeats (450 base-pairs (bp)) and B

repeats (400 bp). Unit one includes one A repeat (A1) and four B repeats (B1 to B4) and unit two consists of one A repeat (A2) and three B repeats (B5 to B7). The human domain displays 7 CTCF binding sites (FREVEL *et al.* 1999) (Fig. 8a). In mice, the *H19* DMD region is arranged in four repeats (A1-4) which carry 4 CTCF binding sites. The sites of CTCF-binding sequences are highly conserved from drosophila to mice and humans (MOON *et al.* 2005). *IGF2* and *H19* share enhancers downstream of *H19* that act on either gene according to parental origin (CHARALAMBOUS *et al.* 2004).

At ICR1, CTCF binds the maternal unmethylated allele and prevents the *IGF2* gene promoter from interacting with enhancers downstream from the *H19* gene, resulting in transcriptional silencing of the maternal *IGF2* gene (MURRELL *et al.* 2004). The methylation of CpG at ICR1 on the paternal allele prevents CTCF binding and allows *IGF2* activation (FILIPPOVA 2008) (Fig. 8a).

Therefore, a proper imprinting of *Igf2* and *H19* requires differential methylation at ICR1 on the parental alleles and CpGs mutation or deletion in this region leads to disruption of *Igf2* and *H19* imprinting (IDERAABDULLAH *et al.* 2008; LI *et al.* 2008; MURRELL *et al.* 2004).

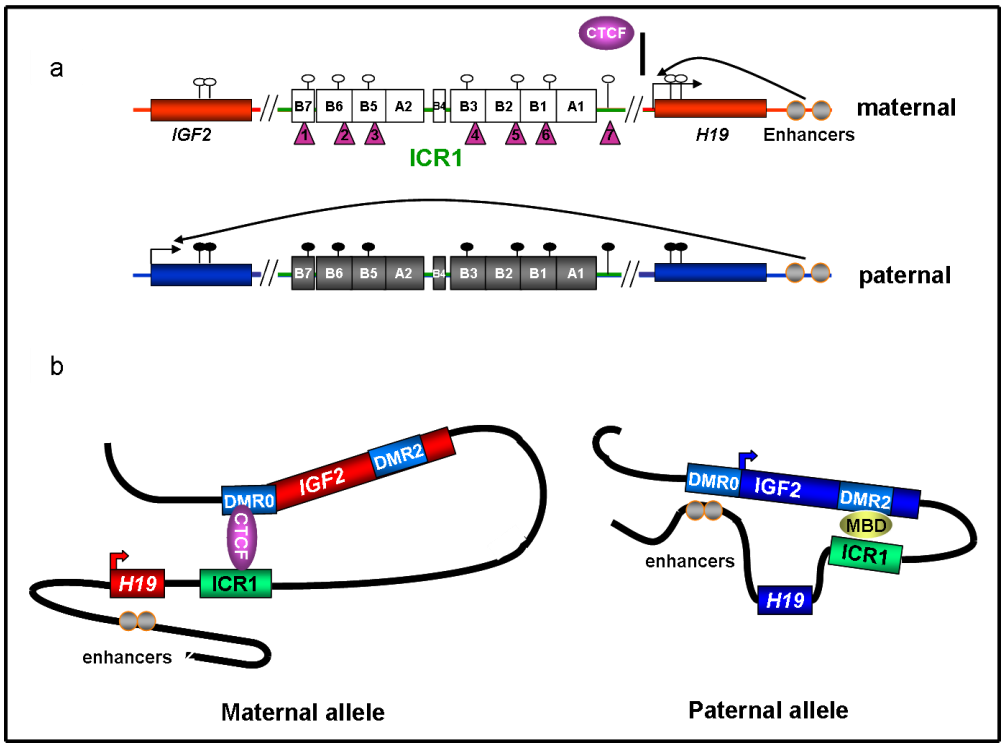


Figure 8: *IGF2/H19*-ICR1 domain and parent-specific interactions. **a**) ICR1 consists of A and B repeats (rectangles). Every domain includes one A repeated sequence and 3 or 4 B repeated sequences (B4 is incomplete). Six of seven CTCF binding sites (violet triangles) are located in the B repeats. Maternal ICR1 is unmethylated (empty circles) and allows CTCF (violet oval) binding at the 7 CTCF binding sites. Methylation on the paternal ICR1 (filled circles) prevents CTCF binding. **b**) ICR1-*IGF2* DMRs parent-specific interactions. On the maternal chromosome, the unmethylated ICR1 binds CTCF and interacts with a region at the 5' end of the *IGF2* gene, DMR0. This loop positions *IGF2* away from the enhancers (which are common for *H19* and *IGF2*) while *H19* is close to the enhancers and suppresses and activates gene expression, respectively. On the paternal chromosome, methylated ICR1 interacts with methylated *IGF2* DMR2. This loop positions *IGF2* close to the enhancers and allows *IGF2* expression (Adapted from (MURRELL *et al.* 2004)).

1.2.4.2 The CTCF protein

The CCCTC-binding factor (CTCF) is a widely expressed zinc finger (ZF) nuclear protein and is known as a versatile transcription factor due to the combinatorial use of its ZFs to bind various DNA sequences (OHLSSON *et al.* 2001). CTCF binds to a number of different sequences which are highly conserved among vertebrates (OHLSSON *et al.* 2001). CTCF is encoded by the *CTCF* gene which maps to the human

chromosome 16q22 and consists of 10 exons (E1-E10). Exons E2-E9 encodes for the 11 highly conserved zinc fingers (FILIPPOVA *et al.* 1996; OHLSSON *et al.* 2001) (Fig. 9a and b).

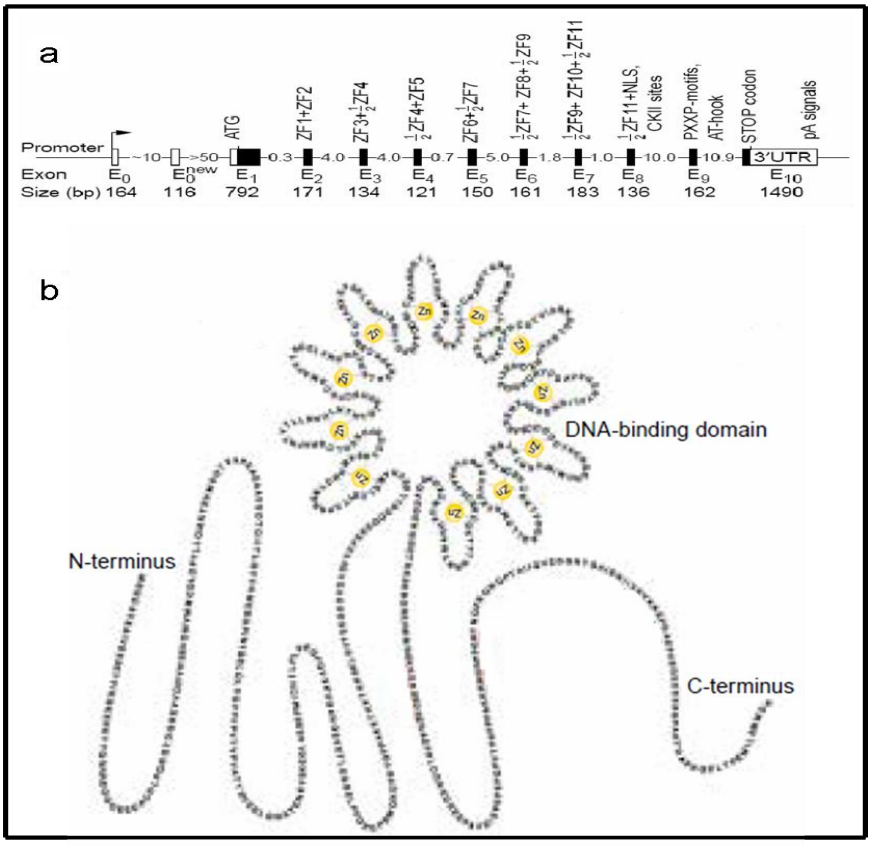


Figure 9: Organization of the human *CTCF* gene (a) and its protein product (b). **a**) Filled boxes, protein coding exons; open boxes, untranslated exons; arrow, transcription start sites. **b**) The human CTCF protein consists of the DNA-binding domain, which is composed of 11 ZFs (OHLSSON *et al.* 2001).

CTCF can regulate gene expression directly by acting as a classical transcriptional factor or indirectly by acting as a chromatin insulator (FILIPPOVA 2008). This indirect regulation is mediated through CTCF ability to form open or closed chromatin loop domains, which allow or block the access of promoters to the nearby enhancers (LI *et al.* 2008; MURRELL *et al.* 2004). The insulator activity of CTCF was first identified in the chicken *β-globin* locus (BELL *et al.* 1999) and later in *H19* DMD (MURRELL *et al.* 2004; OHLSSON *et al.* 2001; PANT *et al.* 2003). Recently by using chromatin conformation capture, the chromatin interactions have been investigated in several studies. CTCF mediates physically

intrachromosomal interactions among *cis*-regulatory elements, within the *IGF2/H19*–ICR1 domain (FILIPPOVA 2008; LI *et al.* 2008). In addition, CTCF mediates interchromosomal interactions, where one allele from one chromosome interacts with another allele on another chromosome. For example, one allele of the mouse *H19* DMD on chromosome 7 interacts with one allele of *Wsb1/Nf1* on mouse chromosome 11 (KRUEGER and OSBORNE 2006; LING *et al.* 2006).

Several models for the epigenetic control of insulators by organising higher order chromatin interactions have been proposed (FILIPPOVA 2008; LI *et al.* 2008; MURRELL *et al.* 2004). At the unmethylated maternal allele, *H19* DMD, which binds the CTCF protein, interacts with *Igf2* DMR1 and *Igf2* promoters and as a result the chromatin is organized in an inactive loop around the *Igf2* gene (LI *et al.* 2008; MURRELL *et al.* 2004). The paternally methylated *H19* DMD, which binds MBD proteins, interacts with the methylated *Igf2* DMR2, resulting in activation *of Igf2* (MURRELL *et al.* 2004). The ICR1 and DMRs interactions could be the same in humans. However, DMR0 instead of DMR1 interacts with ICR1 because DMR1 is absent in humans *IGF2* (MURRELL *et al.* 2008) (Fig. 8b).

The crucial role of CTCF in imprinting regulation has been also identified at other imprinted loci, (but not all of them), such as *KvDMR1* and *Rasgrf1.* This data indicates a prevalent function of the CTCF protein in the regulation of genomic imprinting (FILIPPOVA 2008).

The ICR1 insulator function of the CTCF protein might be altered through two mechanisms: mutation of CTCF binding sites at ICR1 or mutation of the CTCF ZFs (FILIPPOVA 2008). i) Disruption of CTCF function, therefore, can be a result of epigenetic defect at CTCF binding sites and consequently lead to loss of genomic imprinting. Mice studies have shown that mutation of the ICR1 at *H19* DMD locus results in loss of *Igf2* imprinting (ENGEL *et al.* 2004; FILIPPOVA 2008; LI *et al.* 2008; SAKATANI *et al.* 2005) and enhancers tumorigenesis (SAKATANI *et al.* 2005). ii) Alternatively, mutation of the *CTCF* gene might disrupt the

CTCF functions by altering the binding of CTCF ZFs at its target sequences (FILIPPOVA 2008). Only point mutations of the CTCF ZFs have been found in several tumors including Wilms' tumors (FILIPPOVA *et al.* 2002). Interestingly, all these mutations lead only to change in the CTCF function rather than complete loss of CTCF function. Indeed, complete loss of CTCF function is not compatible with life (FILIPPOVA 2008; FILIPPOVA *et al.* 2002).

CTCF function is also regulated by post transcriptional modification (poly(ADP-riposy)lation). There is a functional link between poly(ADP-riposy)lation and CTCF-dependent chromatin insulator function, where disruption of poly(ADP-riposy)lation leads to loss of insulator function not only at the ICR1 locus but also at most other CTCF target sequences (FILIPPOVA 2008; YU *et al.* 2004).

1.2.4.3 Other factors possibly involved in the control of imprinting at ICR1

- MBDs

Methylated CpG binding domain (MBD) proteins are likely candidates for recognizing DNA methylation marks and silencing transcription from imprinted loci. The MBD family includes 5 members, MeCP2 and MBD1 to 4 and these MBD family members bind the methylated DNA with the exception of MeCP2 that binds both methylated and unmethylated DNA, with a preference for methyl-CpG-containing substrates (DHASARATHY and WADE 2008).

Two MBD proteins have been shown to be involved in the regulation of imprinting at ICR1: MeCP2 and MBD3. MeCP2 has been found to bind to the methylated paternal ICR1 allele and other factors to form a repressor complex consequently silencing *H19* expression (DREWELL *et al.* 2002). More recently, Reese *et al.* (REESE *et al.* 2007), has reported that MBD3 binds to the methylated ICR1 allele and its binding seems to

be specific for ICR1 where a depletion of MBD3 reduces the normal DNA methylation at ICR1 but not at other imprinted loci.

- Antisense RNA

At the ICR1 domain, *IGF2* expression is also regulated by noncoding antisense transcript. Antisense transcripts have been proposed to regulate gene expression by *cis*- or *tran*-acting mechanisms (ROYO and CAVAILLE 2008). Recently, a new transcript named *91H* RNA, which is expressed from the maternal allele within the *H19* gene region, has been found to maintain *IGF2* expression in *trans* (on the paternal allele), rather than affecting *H19* expression (BERTEAUX *et al.* 2008).

- Polycomb complex

Polycomb group proteins play a role in silencing imprinted genes through modulation of chromatin structure. EZH2 belongs to one of the two polycomb protein complexes (PRC2/3) and recruits additional polycomb protein complexes, the binding of which contributes to formation of a repressive chromatin state (LEWIS *et al.* 2004; MAGER *et al.* 2003; UMLAUF *et al.* 2004). A recent study by Li *et al*. (LI *et al.* 2008), has suggested that binding of CTCF at ICR1 on the maternal allele recruits polycomb repressive complex 2 (PRC2) and induces H3K27 methylation at *Igf2* promoters. This chromatin complex around the *Igf2* promoters suppresses the expression of the *Igf2* gene on the maternal allele.

1.3 Role of the 11p15 region in development and diseases

Most imprinted genes are associated with foetal and placental growth and aberrant imprinting leads to various growth disorders (ENKLAAR *et al.* 2006; GICQUEL and LE BOUC 2006; GICQUEL *et al.* 2008). This is well illustrated for the 11p15 region which is involved in the aetiology of the Beckwith-Wiedemann (BWS; OMIM 130650) and Silver-Russell (SRS; OMIM 180860) syndromes with foetal overgrowth and foetal growth

retardation, respectively (DELAVAL *et al.* 2006; GICQUEL *et al.* 2005; REIK and WALTER 2001).

1.3.1 Mouse models

To understand the role of the human 11p15 region in development and diseases, several studies have been conducted on the orthologous distal chromosome 7 region in mouse, in particular, the *H19* DMD. Mouse models with transactivation of *Igf2* or knockout of *Cdkn1c* develop foetal overgrowth and some other BWS features (BAKER *et al.* 1993; ENKLAAR *et al.* 2006; GICQUEL and LE BOUC 2006). Mice with combination of these two mutations exhibit more severe BWS phenotypes (ENKLAAR *et al.* 2006). A deletion of the *H19* DMR also results in foetal overgrowth, when maternally transmitted, due to biallelic expression of the *Igf2* gene (DELAVAL *et al.* 2006). Conversely, invalidation of *Igf2* results in foetal growth retardation (BAKER *et al.* 1993; GICQUEL and LE BOUC 2006). Interestingly, point mutations of the CpG in the *H19* DMD result in reduced *Igf2* expression with foetal growth retardation when paternally transmitted. This seems to be caused by disruption of maintenance of DMD methylation and consequently establishment of insulator activity on the paternal allele which prevents *Igf2* expression and allows biallelic *H19* expression (ENGEL *et al.* 2004).

1.3.2 Human models

1.3.2.1 Foetal growth disorders

As described above, the human chromosome 11p15 contains a cluster of imprinted genes which play a crucial role in the control of foetal growth and aberrant genomic imprinting of the 11p15 region has a pivotal role in both BWS and SRS foetal growth disorders (DELAVAL *et al.* 2006; GICQUEL *et al.* 2008; GICQUEL *et al.* 2005; REIK and WALTER 2001).

Beckwith-Wiedemann syndrome is characterized by pre- and/or postnatal overgrowth, macroglossia, abdominal wall defects, organomegaly, body asymmetry, hypoglycaemia in the neonatal period, ear abnormalities and an increased risk of childhood tumours, in particular Wilms' tumor (ENKLAAR *et al.* 2006; GICQUEL *et al.* 2008). The phenotype of BWS patients is variable and children with BWS do not display all these features together (ENKLAAR *et al.* 2006; GICQUEL *et al.* 2008).

In contrast to BWS phenotypes, SRS is characterized by intra-uterine and postnatal growth retardation with spared cranial growth and body asymmetry. Other clinical features include a distinctive triangular face with prominent forehead and a pointed chin, clinodactyly of the fifth fingers, severe feeding difficulties, cafe´-au-lait spots, genital abnormalities and hypoglycaemia (ROSSIGNOL *et al.* 2008).

1.3.2.2 Genetics of foetal growth disorders

Unlike most other genetic syndromes, BWS and SRS can be caused by different genetic and epigenetic abnormalities, which lead to abnormal imprinting (ENKLAAR *et al.* 2006; GICQUEL *et al.* 2008) (Table II). Epigenetic defects are much more frequent in 11p15 growth disorders than other imprinting disorders and account for more than 60% of BWS and SRS cases. This suggests that the 11p15 region is particularly vulnerable to epigenetic cues (GICQUEL *et al.* 2008; ROSSIGNOL *et al.* 2008). Loss of DNA methylation involving the ICR2/*KCNQ1* domain results in 60% of BWS cases whereas gain of DNA methylation involving the *IGF2/H19*-ICR1 domain results in 10% of BWS cases (GASTON *et al.* 2001). Patients with gain of methylation at ICR1 have a higher risk of developing tumors particularly Wilms' tumor than patients belonging to other molecular subgroups (COOPER *et al.* 2005; GASTON *et al.* 2001). Gain of DNA methylation at ICR1 results in switch from the

maternal to the paternal epigenotype and consequently, biallelic expression of *IGF2* and silencing of *H19* (ENKLAAR *et al.* 2006).

Maternal uniparental disomy (mUPD) for chromosome 7 accounts for approximately 10% of SRS cases and abnormalities of the 11p15 region account for about 60 % of SRS cases (table II). Most 11p15 SRS cases display epigenetic alterations (loss of ICR1 methylation on the paternal allele) and rare cases display maternal 11p15 duplication (ROSSIGNOL *et al.* 2008). Gicquel *et al*. (GICQUEL *et al.* 2005) reported the first cases of SRS with partial loss of methylation on the paternal ICR1 and *H19* promoter. This switch from the paternal to the maternal epigenotype results in biallelic expression of the *H19* gene and downregulation of the *IGF2* expression (GICQUEL *et al.* 2005).

Table II: Frequencies of 11p15 genetic and epigenetic defects in BWS and SRS disorders

Molecular Defect	BWS		SRS	
	Parental origin	Frequency	Parental origin	Frequency
11p15 UPD	paternal UPD	20%		
11p15 duplication	paternal	<2%	Maternal	<4%
ICR1 epimutation	gain of methylation	10%	loss of methylation	60%
ICR2 epimutation	loss of methylation	60%		
CDKN1C mutation	maternal	5%		

UPD: uniparental disomy

Whatever the methylation defect is, in either BWS or SRS, the DNA methylation defect has a mosaic distribution suggesting that the defect is a post-zygotic event (ROSSIGNOL *et al.* 2008). In agreement with this suggestion several studies reported a high incidence of discordant monozygotic twins in both BWS and SRS patients. Mosaicism could explain the high variability of the clinical phenotype among individuals in both BWS and SRS (ROSSIGNOL *et al.* 2008).

Although the human 11p15 DNA methylation defects in BWS and SRS caused by abnormal DNA methylation at ICR1 have been investigated in several studies, the mechanism(s) resulting in gain or loss of DNA methylation is still unknown in most cases. In a few familial BWS cases with a gain of methylation at ICR1, deletions within ICR1 have been reported (fig. 10); deletions result in a BWS phenotype only if the deletion is maternally-inherited (CERRATO *et al.* 2008; PRAWITT *et al.* 2005; SPARAGO *et al.* 2004; SPARAGO *et al.* 2007). Subjects who inherit the deletion from the father do not display any phenotype and, more particularly, no SRS phenotype. These deletions (1.4 to 2.2 kb) removed from one to three CTCF binding sites and in most BWS cases, result in partial gain of methylation on the maternal allele and disturbance of genomic imprinting at the *H19/IGF2*-ICR1 locus (CERRATO *et al.* 2008; PRAWITT *et al.* 2005; SPARAGO *et al.* 2004; SPARAGO *et al.* 2007). The 2.2 kb deletion, which abolished 3 CTCF binding sites, did not result in gain of methylation (PRAWITT *et al.* 2005). It was proposed at this time that alteration of the spacing rather than loss of CTCF binding sites might be essential for ICR1 gain of methylation (PRAWITT *et al.* 2005). Very recently, Scott *et al*. (SCOTT *et al.* 2008) showed that a larger deletion of 5.3 kb removing both repeat blocks and six of the seven CTCF binding sites also results in gain of methylation at the seventh CTCF binding site. Interestingly, a 2.9 kb insertion adding two CTCF binding sites between the two ICR1 repeat blocks was also shown to result in gain of methylation (fig. 10). Although the insertion did not result in loss of CTCF binding sites or disturbance of the architecture within each block, there was a gain of methylation at ICR1 (SCOTT *et al.* 2008).

Several attempts have been made to identify human ICR1 mutations in SRS cases but, neither mutations nor deletions have been found at the ICR1 locus (including 3 familial cases) (BARTHOLDI *et al.* 2008; BLIEK *et al.* 2006; BRUCE *et al.* 2008; YAMAZAWA *et al.* 2008).

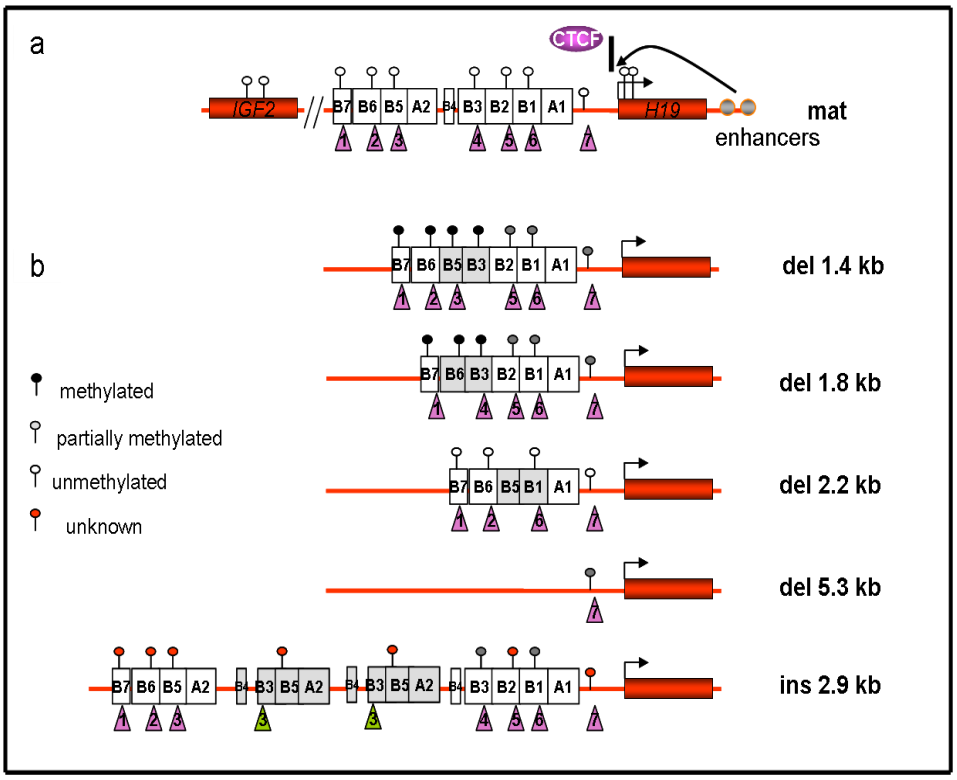


Figure 10: Schematic diagram represents wild-type and mutations in humans *IGF2/H19*-ICR1 domain. **a**) Wild-type ICR1. **b**) Consequences of different deletions on the structure of ICR1 demonstrated in humans. Four deletions at ICR1 which abolish one CTCF binding site in 1.4 kb deletion (SPARAGO *et al.* 2004) or many CTCF binding sites (two sites in 1.8 kb deletion (PRAWITT *et al.* 2005), three sites in 2.2 kb deletion (SPARAGO *et al.* 2007) and six sites in 5.3 kb deletion (SCOTT *et al.* 2008). The deletions of 1.4 kb and 1.8 kb resulted in fusion of two repetitive domains and they were associated with a gain of methylation at ICR1. The 2.2 kb deletion resulted in an allele that was similar to one of the two repetitive units in terms of number of repeated sequences and the spacing of CTCF binding sites. This deletion (2.2 kb) did not cause any change in methylation status at ICR1. In contrast, there was an insertion (2.9) (SCOTT *et al.* 2008) that rearranges the repetitive units at ICR1. This insertion added two CTCF binding sites by fusion of one A and parts of two B repeats (B3 and B5) and it was association with a gain of methylated at ICR1.

1.4 Aims of the research

Increasing evidence points towards the importance of genomic imprinting in development and diseases. The differentially methylated ICR1 mediates the reciprocal expression of the imprinted genes (*IGF2* and *H19*) and the zinc finger CTCF protein is required for the maintenance of differential methylation on the maternal allele.

Epigenetic defects (gain or loss of DNA methylation) of the human ICR1 11p15 domain result in two opposite foetal growth disorders (BWS and SRS) depending on which parental allele is affected. A few deletions within ICR1 have been reported in familial BWS cases with ICR1 gain of methylation, however the mechanism(s) of the DNA methylation defects at ICR1 remains largely unknown in most BWS and SRS patients. The aim of this project was to extensively characterize the ICR1 domain and the *CTCF* gene in a series of BWS and SRS patients with gain and loss of methylation at ICR1, respectively.

2 SUBJECTS AND METHODS

2 SUBJECTS AND METHODS

2.1 Subjects

This project was conducted in compliance with institutional guidelines for research studies in human genetics and informed consent was obtained from participating individuals and/or their parents. The study population consisted of 37 patients with growth disorders caused by a DNA methylation defect of the 11p15 ICR1 domain. Blood DNA was available for all patients and DNA extracted from tongue tissue was available for two patients. The patients had already been diagnosed with a DNA methylation defects at ICR1 by Methyl-sensitive southern blotting of blood DNA.

Twenty-one patients were diagnosed with BWS.

BWS was diagnosed in all patients according to the clinical signs described in table III. The molecular investigation was performed in these patients between the age of 0 and 17 years. Analysis of the 11p15 region by methyl-sensitive southern blotting showed a gain of methylation at ICR1 (CTCF binding site 3) and/or the *H19* promoter and a normal methylation pattern at ICR2 in all BWS patients. One of these patients was diagnosed during foetal life and the parents decided to have the pregnancy terminated.

Two BWS patients were siblings and the family pedigree is shown in Fig. 15a. Patient II-1 was born after 39 weeks of gestation from non consanguineous parents. The clinical presentation at birth included macrosomia, severe macroglossia, diastasis recti, nephromegaly, neonatal hypoglycemia and bilateral cryptorchidism. A sibling (II-3) displaying macrosomia, macroglossia and nephromegaly died at birth from unknown reason. Another brother (II-5) displayed the same phenotype as patient II-1 including macrosomia and nephromegaly. Both patients (II-1 and II-5) underwent partial glossectomy within the first months of life. Their father I-1 was born with severe macrosomia

but did not display any other BWS phenotypes except a supernumerary nipple. There was no familial history evocative of BWS on the father's side (seven siblings). The two sisters (II-2 and II-4) as well as the mother (I-2) were born with a normal birth weight and were phenotypically normal. The pedigree was indicative of an X linked disorder but sequencing of the *Glypican 3* gene ruled out a Simpson-Golabi Behmel syndrome. Analysis of the 11p15 region showed an isolated gain of methylation at ICR1 and the *H19* promoter in patients II-1 and II-5 and a normal methylation pattern in one of the sisters and the two parents. Seven of 20 (35%) BWS patients developed a Wilms' tumor (median age: 24 months; range: 14-39 months).

Sixteen patients were diagnosed with SRS

All SRS patients had a severe phenotype with severe growth retardation, relative macrocephaly at birth, facial dysmorphy and feeding difficulties; most of them displayed body asymmetry (table III). Molecular analysis was performed at the age of 0 to 7 years. Analysis of the 11p15 region showed a loss of methylation at ICR1 (CTCF binding site 3) and/or *H19* promoter and a normal methylation status at ICR2. Two SRS patients were monozygotic twins and both of them had a co-twin who was unaffected (monozygotic discordant twining).

Table III: Clinical features of patients with BWS and SRS

BWS		**SRS**	
phenotype	BWS patients	phenotype	SRS patients
n	21	n	16
sex F/M	11/10	sex F/M	7/9
Macrosomia, n (%)	19/21 (90.5)	IUGR, n (%)	16/16 (100)
Macroglossia, n (%)	18/21 (86)	post-natal growth retardation, n (%)	16/16 (100)
abdominal wall defect, n (%)	13/21 (62)	relative macrocephaly at birth, n (%)	16/16 (100)
diastasis recti, n (%)	9	facial dysmorphy, n (%)	16/16 (100)
umbilical hernia, n (%)	4	body asymmetry, n (%)	13/16 (81)
Exomphalos, n (%)	0	feeding difficulties, n (%)	12/14 (86)
organomegaly, n (%)	16/18 (89)	developmental delay, n (%)	3/14 (21)
body asymmetry, n (%)	7/21 (33)	clinodactyly, n (%)	10/13 (77)
hypoglycemia, n (%)	5/19 (26)		
Wilms' tumor, n (%)	7/20 (35)		

IUGR: intrauterine growth retardation

2.2 Methods

2.2.1 DNA extraction

2.2.1.1 DNA extraction from blood tissues

DNA was extracted from blood samples as described previously (MILLER *et al.* 1988). Briefly, blood samples (5-10 ml) were transferred in 50 ml Falcon tubes with 40 ml lysis buffer (20 mM TRIS pH 7.5 and 5 mM MgCl2), incubated on ice for 15 min and centrifuged at 3000 rpm for 5 min at 4°C. Supernatant was discarded and pellets were resuspended in 30 ml lysis buffer followed by centrifugation at 3000 rpm for 5 min at 4°C. The pellets were resuspended in 3.5 ml protein lysis buffer (20 mM TRIS pH 7.5, 400 mM NaCl and 2 mM EDTA). The samples were digested overnight at 56°C with 200 µl of SDS 10% and 80 µl of proteinase K (10 mg/ml). After digestion, 1 ml of saturated NaCl (6 M) was added to each tube and shaken for 15 sec, followed by centrifugation for 15 min at 4000 rpm. Supernatant containing the DNA was transferred to another 15 ml tube. Two volumes of absolute ethanol were added and the tubes were inverted several times until the DNA precipitated. The precipitated DNA was transferred to a 1.5 ml eppendorf tube and diluted in TE buffer (10 mM TRIS pH 7.5 and 1 mM EDTA) at a concentration of 1 µg/ µl.

2.2.1.2 DNA extraction from tongue tissues

Three to four slices (~1 mm each) of tongue tissue were washed twice with 1 ml of PBS. The tissue was resuspended in 1 ml of the following buffer (9 mM TRIS pH 8, 26 mM EDTA and 0.5% SDS) and 10 µl of proteinase K (20 mg/ml) and shaken gently. The mix was incubated 3 hrs at 56°C then overnight at 37°C. The following day, 430 µl of saturated NaCl was added to the solution and mixed gently before centrifuged at 12000 rpm for 20 min at 4°C. The supernatant was transferred to a tube containing 3 ml of absolute ethanol to precipitate the DNA. The precipitated DNA was transferred to a 1.5 ml tube with 1 ml of TE buffer and rotated for 12 hrs at room temperature.

2.2.2 Polymerase chain reaction (PCR)

2.2.2.1 The IGF2/H19-ICR1 region

The ICR1 region is a very CG rich region which is particularly difficult to amplify. Therefore, different polymerase enzymes, different protocols of PCR and different enhancers have been investigated in this study.

The primer sets used to amplify the target sequences to analyze ICR1 and the CTCF binding sites 1 to 7, and the PCR conditions are summarized in table IV. Each PCR reaction mix contained 0.4 µM of each primer (forward and reverse), 0.3 mM of deoxynucleotides (dNTPs), 60-100 ng DNA and different concentrations of polymerase enzymes and PCR reagent buffers depending on the of primer sets used. The reactions were amplified using the Gene Amp PCR system 9700 (Applied Biosystems). To investigate deletions in the ICR1 region, Long Range PCR amplifications were performed using the BIO-X-ACT™ Long DNA Polymerase (BIOLINE, Boston, MA, USA) and the primers used by Scott *et al.* (SCOTT *et al.* 2008). Additional primers were designed along the ICR1 domain using Primer3 software (http://fokker.wi.mit.edu/primer3/input.htm) to finely characterize the deletion (table IV).

The seven CTCF binding sites were investigated using primers previously described by Bliek *et al*. (BLIEK *et al.* 2006). CTCF binding sites from 1 to 3 were amplified by one PCR using GC-RICH Enzyme Mix (Roche Applied Science, Australia). CTCF binding sites 4 to 6 region was more difficult to amplify thus different enhancers and GC-RICH resolution solution (provided in the kit) concentrations (from 0.5 to 2.5 M) were investigated. The best amplification was obtained by using 1 M of GC-RICH resolution solution. “Touch-down” PCR also helped to reduce the amplification of non-specific sequences. Standard PCR condition was conducted to amplify the CTCF binding site 7 using FailSafe DNA polymerase (EPICENTRE Biotechnologies, Australia).

2.2.2.2 The CTCF gene

Standard PCR amplifications were performed to analyse all exons (1 to 10) and flanking intronic regions of the *CTCF* gene using AmpliTaq Gold DNA polymerase (Applied Biosystems). Most primers used to amplify each target sequence have been previously described (CUI *et al.* 2001; YEH *et al.* 2002) and some additional primers have been designed using Primer3 software (table V). Each PCR reaction mix contained 0.4 µM of each primer (Forward and Reverse), 0.3 mM of dNTPs, 1.5 mM of MgCl2 Buffer and 50-80 ng DNA.

2.2.3 Gel electrophoresis and DNA purification

Before purification, the PCR products of CTCF binding sites 1-3 and 4-6 were separated on a 0.8 agarose gel at 100 volt (V) and the correct size fragment was cut from the gel. Small amount of CTCF binding site 7 and *CTCF* exons PCR products were also electrophoresed in 0.8 agarose gel at 100 V to check the PCR products and remaining PCR products were purified. The gel and PCR product purifications were performed using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA) according to the manufacturer's instructions.

2.2.4 Sequencing and sequencing analysis

Different forward and reverse primers were used in the sequencing to analyse the different amplified regions (table IV and V). Sequencing was performed by the sequencing lab at the Baker IDI Heart and Diabetic Institute (Melbourne, Australia) using the BigDye Terminator V3.1 cycle sequencing kit on the ABI 3100 automatic sequencer (Applied Biosystems, Foster City, CA, USA). The sequencing results were analyzed with the Chromas 1.45 software (Applied Biosystems, Foster City, CA, USA).

2.2.5 Bisulfite treatment of DNA

DNA methylation patterns at ICR1 in BWS and SRS patients with ICR1 mutations and in control samples were analysed by bisulfite sequencing. Treatment of DNA with bisulfite results in converting all unmethylated cytosines residues to uracils while 5-methylcytosine residues remain unaltered. Bisulfite treatment thus gives different DNA sequences depending on wether the cytosine of a CpG nucleotide is originally methylated or unmethylated.

Genomic DNA was bisulfite modified using the EZ DNA modification kit (Zymo Research, Orange, CA, USA) according to the manufacturer's instructions. Briefly, 5 µl of M-Dilution buffer were added to 1 µg of genomic DNA from blood or tongue tissues in a total volume of 50 µl. After incubation of the reaction mix at 37°C for 15 min, 100 µl of CT conversion reagent (EZ DNA modification kit) was added to each sample and incubated in the dark for 16 hours (hrs) at 50°C. Bisulfite treated DNA was washed using reagents provided in the kit and then desulphonated by adding 200 µl M-desulphonation Buffer (EZ DNA modification kit). Modified DNA was washed twice and then eluted in 20 µl H2O.

2.2.6 Amplification of bisulfite-treated DNA

After treatment of samples by bisulfite, specific primer sets were used to amplify each CTCF binding site sequence (1 to 7) (table VI). Each primer set was designed to be bisulfite-specific and to amplify both methylated and unmethylated alleles. Therefore, all the cytosines were converted to thymines in the primers such that they did not amplify non-bisulfite treated DNA. In addition, the primers did not include CpG in their sequences so they amplified both methylated and unmethylated sequences. Some PCR primers were designed using the MethPrimer web based design program (http://www.urogene.org/methprimer/) and others were previously reported (table VI). Standard PCR amplification

was performed using 1 U HotStar DNA polymerase (Qiagen, Australia), 0.4 µM of each primer (forward and reverse), 0.3 mM of dNTP, 1 X buffer (provided in the kit) and 2 µl of bisulfite treated DNA. Following the size confirmation of PCR fragments by agarose gel, the fragments were cloned and sequenced.

2.2.7 Cloning and sequencing of PCR products

To characterize the 1.8 kb deletion precisely, the PCR product from PCR with primers ICR1 _4426_5166_F (4426 F) and ICR1_6135_6969_R (6969 R) was cloned and sequenced. Cloning was also conducted after treatment of the DNA by bisulfite to look at individual methylation status from one allele of one cell.

After size confirmation of the fragments of different PCR products by electrophoresis, the PCR products were cloned using TOPO TA Cloning® Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations. In brief, 2 µl from each PCR product was cloned into a pCR®II-TOPO vector and incubated for 20 min at room temperature. Two µl of the above mixture was added into a vial of 'One Shot® DH5α™-T1R Competent Cells and incubated on ice for 30 min before transformation. The cells were heat-shocked at 42°C for 45 seconds (sec). Following Transformation, the cells were placed on ice and 250 µl of S.O.C media (provided in the kit) was added then the mixtures were shaken at 250 rpm for 1 hr at 37°C. Fifty to sixty µl of each transformation was spread on LB plate containing 50 µg/ml kanamycin and 40 mg/ml of X-gal (for blue/white screening) and the plates were incubated at 37°C overnight. The following day, 10-12 white single bacterial colonies per sample was transferred to 15 ml tubes containing 3 ml of LB medium with 50 µg/ml kanamycin. Culture media were then incubated at 37°C overnight. The recombinant plasmid DNA from each colony was purified for each sample using the Wizard® Plus SV Minipreps DNA purification System kit (Promega, Madison, WI, USA) as described in the kit protocol.

Inserted DNA in the plasmids was sequenced with M13 Forward (-20) or M13 Reverse primer (provided in the kit). Sequencing was performed by the sequencing laboratory at the Baker IDI Heart and Diabetic Institute (Melbourne, Australia). The sequencing results were analyzed with the Chromas 1.45 software (Applied Biosystems, Foster City, CA, USA).

Analysis of the methylation pattern for all DNA fragments was performed using QUMA Quantification tool (http://quma.cdb.riken.jp/). Each CTCF binding site sequence for each patient was compared to a non-treated DNA sequence and the percentage of methylation was estimated using the following formula:

Methylation index (%) = number of methylated CpG sites/number of total CpG sites X 100

Table IV: Primers and PCRs conditions used for analysis of the human ICR1 region

Primers				PCR				Primers References
Name	**Sequence**	**Position ***	**Purpose**	**Polymerase**	**Buffer &/or enhancer**	**Cycles**	**Annealing temp**	
ICR1 scott H1 F	5'-CACATCTGCCTCCAGGTGAC-3'	2704	amplification					(Scott *et al.* 2008)
ICR1 scott H1 R	5'-GAGTGTCAAAGCCGTGAAGG-3'	9241						
ICR1_2819_3597_F**	5'-ACCCAGGCACTCACCATTAC-3'	2819						
ICR1_2819_3597_R**	5'-CCAAATTTCCTGGGAGACTG-3'	3597						
ICR1_3414_4182_F**	5'-TCACCCCCAGGAGAGTAGTG-3'	3414						
ICR1_3414_4182_R**	5'-CCAGCCTGGATGATAAGAGC-3'	4182						
ICR1_3981_4685_F**	5'-CCTGGGAAAGGGCTGTTATT-3'	3981	Amplification to	BIO-X-ACT™	1X Hi-Spec			
ICR1_3981_4685_R**	5'-GTTCTTTGGGGTCCAAGTCA-3'	4685	characterize the	Long Mix (2U)	1XOptiBuffer	40	58°C	
ICR1_4426_5166_F**	5'-CCGGCCGATTTTCTGTAATA-3'	4426	1.8 kb deletion	BIOLINE	2 mM MgCl2			
ICR1_4426_5166_R**	5'-CCGAGATCCCTATCATCCAA-3'	5166	in patient KL41					
ICR1_5662_6459_F**	5'-TTCCCCTTCTGTCTCACCAC-3'	5662						
ICR1_5662_6459_R**	5'-CTGATTCCAGCAGCACAGAG-3'	6459						
ICR1_6135_6969_F**	5'-TCAGTGCAGGTTTGAGATGC-3'	6135						
ICR1_6135_6969_R**	5'-CACTTCACTGTCCCCCAAGT-3'	6969						
ICR1_7892_8782_F**	5'-GAGGCTTCTCCTTCGGTCTC-3'	7892						
ICR1_7892_8782_R**	5'-CCGGAAATACAAATGCTCCA-3'	8782						
CTCFSITE 1-3 F	5'-GCCCATCTTGCTGACCTCAC-3'	4825	Amplification	GC-rich (2U)	1.5M GC-RICH	40	62 °C	(Bliek *et al.* 2006)
CTCF SITE 1-3 R	5'-AGAAGACCTCCGAGAACCCTG-3'	5834	& sequencing	Roche				
CTCF SITE 2 R	5'-AATGTGGCTCCCATGAGTG-3'	5461	sequencing					(Bliek *et al.* 2006)
CTCF SITE 4-6 F	5'-GGTAGGACCCTTGTACGAGCC-3'	6864	amplification	GC-rich (2U)	1.5 M GC-RICH	10	72 °C ***	(Bliek *et al.* 2006)
CTCF SITE 4-6 R	5'-GACCTGAAGATCTGGTGCGG-3'	8121	& sequencing	Roche	1 M GC-RICH resolution	40	62 °C	
CTCF SITE 5 R	5'-AGAAGGGTTTCACACTAGGGCCG-3'	7829	sequencing					(Bliek *et al.* 2006)
CTCF SITE 7 F	5'-ATTTCCTGAGTCTCCCCTTGG-3'	8785	Amplification	FailSafe (1.25) EPICENTRE Biotechnologies	1X buffer D	50	60 °C	(Bliek *et al.* 2006)
CTCF SITE 7 R	5'-TCGGCAAACCCTCTGTTCC-3'	9137	& sequencing					

*: According to the reference sequence AF 125183
**: Primers designed using Primer3 software (http://fokker.wi.mit.edu/primer3/input.htm)
***: Touch down PCR 1 °C / cycle for the 10 first cycles

Table V: Primers and PCRs conditions used for the analysis of the human *CTCF* gene

Primers				PCR				Primers References
Name	**Sequence**	**Fragment size**	**Purpose**	**Polymerase**	**Buffer &/or enhancer**	**Cycles**	**Annealing temp**	
CTCF1 YehF	5'-TAAATATGGAGGAACAGCC -3'		sequencing					(YEH *et al.* 2002)
CTCF1 CuiF CTCF1DR	5'-CATCAAGAGCACATGTCTGTTGTG -3' 5'-TGCACTGTGTTGTATGCTTATCC -3'	1007bp	amplification & sequencing	AmliTaq Gold Polymerase (2U)	1X green buffer & 1.5 mM MgCl2	40	60 °C	(CUI *et al.* 2001)
CTCF2 F* CTCF2 R*	5'-CACTTTGAAACTCTGCAGCAA -3' 5'-TCCCGCTGGAGTCAGCTT -3'	300bp					60 °C	
CTCF3 F* CTCF3 R*	5'-TGCCACACATTGAACTCTGTC-3' 5'-TCATGGTCTGCCTAAGAGAGAT-3'	252bp					58 °C	
CTCF4 CuiF CTCF4 CuiR	5'-GCTTTTGTGCCTAACCTACTGTGC-3' 5'-CTGAACAACGAATTCAGAGGATATGC-3'	368bp					60 °C	(CUI *et al.* 2001)
CTCF5 CuiF CTCF5 CuiR	5'-TCTCTGTGGTGTAGCTATTCTG-3' 5'-TGTTATGAGAGTCAGAAGGTGAAGT-3'	350bp					60 °C	(CUI *et al.* 2001)
CTCF6 F* CTCF6 R*	5'- TTCACATTACCCTGGGCTTT-3' 5'-ACCGAGAAAGCACCAACAAC -3'	234bp					58 °C	
CTCF7 CuiF CTCF7 CuiR	5'-CGTGTGGAGTCTAGACCTAGCTTGG-3' 5'-CCATGCTCTGCAGAGGAAGAC-3'	377bp					58 °C	(CUI *et al.* 2001)
CTCF8F* CTCF8R*	5'-CTTCCAATCTGATCTTAGCTTTTT -3' 5'-ACTGGAAACTGGGCTCCAA -3'	255bp					58 °C	
CTCF9 CuiF CTCF9 CuiR	5'-TTCATCTTCCACCACCCTTCTC-3' 5'-GACTTCCTCAGATGTTCCTCAGT-3'	405bp					60 °C	(CUI *et al.* 2001)
CTC10 YehF CTC10 YehR	5'-ATTCTTGGGGCTTTAATGGAC-3' 5'-ATGCGGGCCGTTTAAACACAG-3'	274bp					60 °C	(YEH *et al.* 2002)

*: Primers designed using Primer3 software (http://fokker.wi.mit.edu/primer3/input.htm)

Table VI: Primers and PCRs conditions used for the analysis of the human ICR1 region following bisulfite treatment.

Primers				PCR			Primers references
Name	**Sequence**	**Position***	**Purpose**	**Polymerase**	**Cycles**	**Annealing temp**	
TAKAICTS1F TAKAICTS1R	5'-GTATTTTTGGAGGTTTTTTATTTAG-3' 5'-ACACCTAACCTAAAAAACCTAAAAC-3'	4773 5004	amplification & sequencing	HotStar (1U) Qiagen	38	50 °C	(TAKAI *et al.* 2001)
TAKAICTS2F TAKAICTS2R	5'-AGGTGTTTTAGTTTTTTGGATGATA-3' 5'-CCATAAATATTCTATCCCTCACTA-3'	5131 5451					
TAKAICTS3F TAKAICTS3R	5'-GGTTTTTGGTAGGTATAGAAATTG-3' 5'-CACCTAACTTAAATAACCCAAAAC-3'	5593 5811					
TAKAICTS4F BIS CBS4 R 7302**	5'-GTTTTTGGTAGGTTTAAGAG-3' 5'-CACCCTAAAATAAATCAAACACA-3'	7026 7302					
TAKAICTS5F TAKAICTS5R	5'-TTTTGTAGGGTTTTTGGTAG-3' 5'-TCCCATAAATATCCTATACCTC-3'	7424 7693				55 °C	
ICR1-5962BT ICR1-6413BT	5'-TGTTGAAGGTTGGGGAGATGGGA-3' 5'-CCCAAACCATAACACTAAAACCCTC-3'	7734 8185					(ULANER *et al.* 2003)
CTCFBS7CG2009F** CTCFBS7CG2009R **	5'-TTTTTTATAATGTTTGGAGTAGGAG-3' 5'-TCTCAAACTTTTCCATAAATAAACC-3'						

*: According to the reference sequence AF 125183
**: Primers designed using the MethPrimer web based design program (http://www.urogene.org/methprimer/)

3 RESULTS

3 RESULTS

3.1 Analysis of the *IGF2/H19*-ICR1 domain

3.1.1 Identification of mutations in the ICR1 domain

3.1.1.1 Long range PCR

Long range PCR was conducted on 21 BWS and 16 SRS patients with primers spanning the whole ICR1 (ICR1 Scott H1 F and R) which allowed to identification of a deletion in one BWS case (KL41) (Fig. 11and 12). Analysis of the PCR amplification products on an agarose gel highlighted the deletion of approximately 2000 bp in this patient (Fig. 12). Because it was difficult to sequence the complete allele carrying the deletion (~4500 bp), additional primers were used to localize more precisely the deletion. From PCRs with different primers (Fig. 12b), it was clear that the deletion started between 4426 and 5662 and ended between 6459 and 6969 (according to the reference sequence AF 125183). Then, a PCR was performed using 4426 F and 6969 R primers, which gave the shortest fragment with both alleles, to identify the exact start and end of the deletion. By cloning and sequencing this PCR product (including the normal and deleted alleles), it was found that the deletion started between 5067 and 5086 and ended between 6900 and 6919 (minimum size 1814 bp and maximum size 1852 bp) (Fig. 12c). This deletion was generated by recombination of the B6 and the B3 repeats, removing almost the entire B6 repeat and also the B5, B4 and A2 repeats. Recombination of the B6 with the B3 repeats results in the loss of CTCF binding sites 2 and 3 (Fig. 12c). Parental analysis showed that this deletion was maternally-inherited.

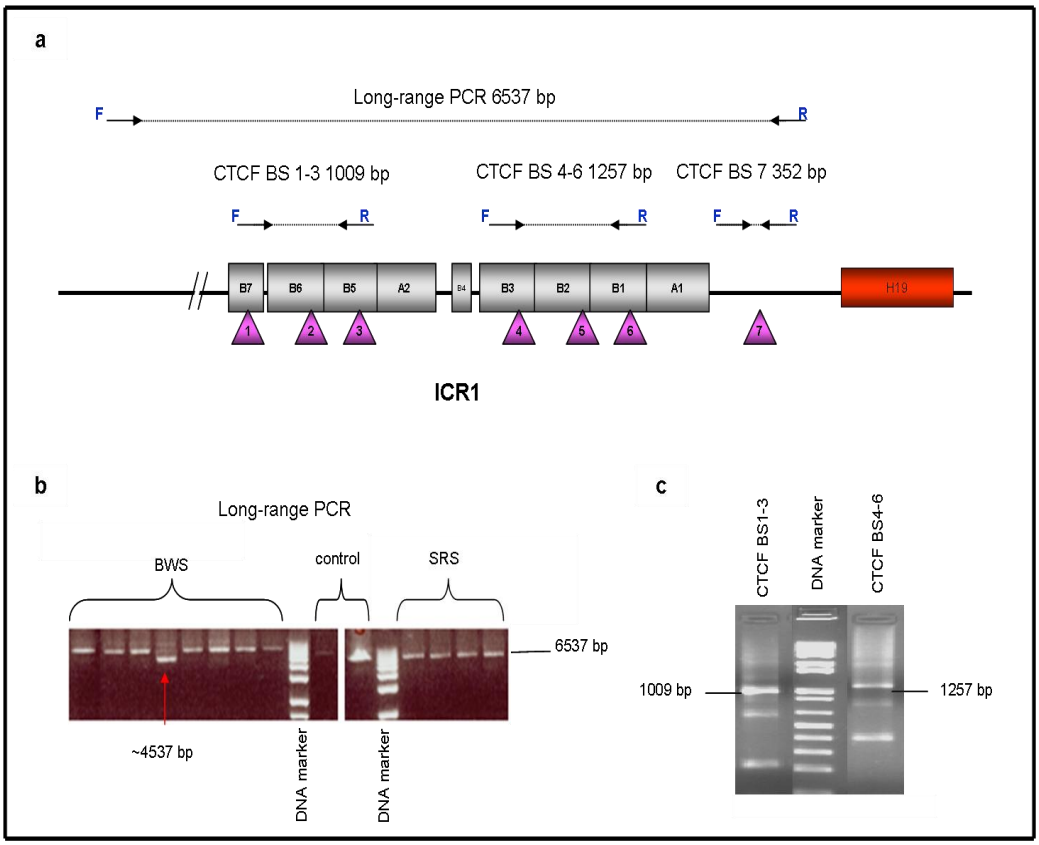


Figure 11: Analysis of the ICR1 domain. **a**) Primer sets for long range PCR and CTCF binding sites (CTCF BSs). Forward primers (F) and reverse primers (R). **b**) Agarose gel of long range PCR products (6537 bp) for BWS and SRS patients. One BWS displayed a deletion, indicated by red arrow, **c**) Agarose gel of PCR products of CTCF BS1-3 and CTCF BS4-6. The fragments 1009 bp of the CTCF BS1-3 and 1257 bp of the CTCF bS 4-6 were cut, purified and sequenced in all patients and control samples (All numbers are numbered according to AF125183).

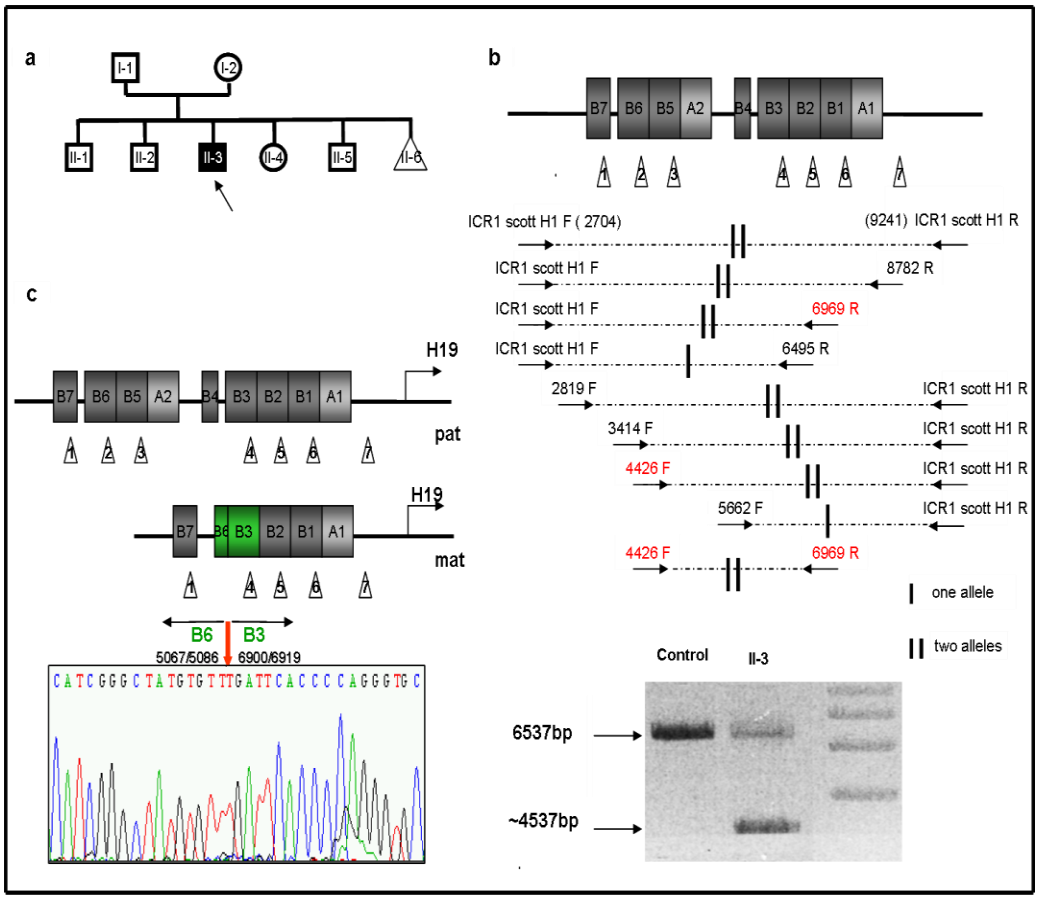


Figure 12: ICR1 analysis in BWS case KL41. **a**) Pedigree of the family. **b**) Identification and characterization of the deletion. Identification of the deletion, using ICR1 scott H1F and R primers, is showed on a 0.8% agarose gel. The 6537 bp band corresponds to the normal allele whereas the ~4537 bp band corresponds to the deleted allele. Additional primers were used to characterize the break points of the deletion. The start and the end of the deletion were identified by the primers which are represented in red colour. **c**) Localization of the deletion: Position of the breakpoints (indicated by arrows at 5067/5086-6900/6919) has been identified after cloning and sequencing both alleles and is highlighted on the chromatogram. The deletion deletes two CTCF binding sites (2 and 3) (All numbers are numbered according to AF125183).

3.1.1.2 Sequencing analysis of ICR1

Direct sequencing of the ICR1 domain was performed for all BWS and SRS patients using 3 PCR conditions for each primer sets, CTCF BS1-3 F and R (1009 bp), CTCF BS4-6 F and R (1257 bp) and CTCF BS7 F and R (352 bp) (Fig. 11a and c). Sequencing analysis of these regions allowed to identify a very small deletion in one BWS patient (CF32), that was not found in control samples (n=50). This deletion was 8 bp long (residues 7277-7284 according to the reference sequence

AF125183) within the B3 repeat. This deletion was outside a CTCF binding site, 116 bases 3' of the fourth CTCF binding site and did not remove any CTCF binding site (Fig. 13). The 8 bp deletion was maternally-inherited.

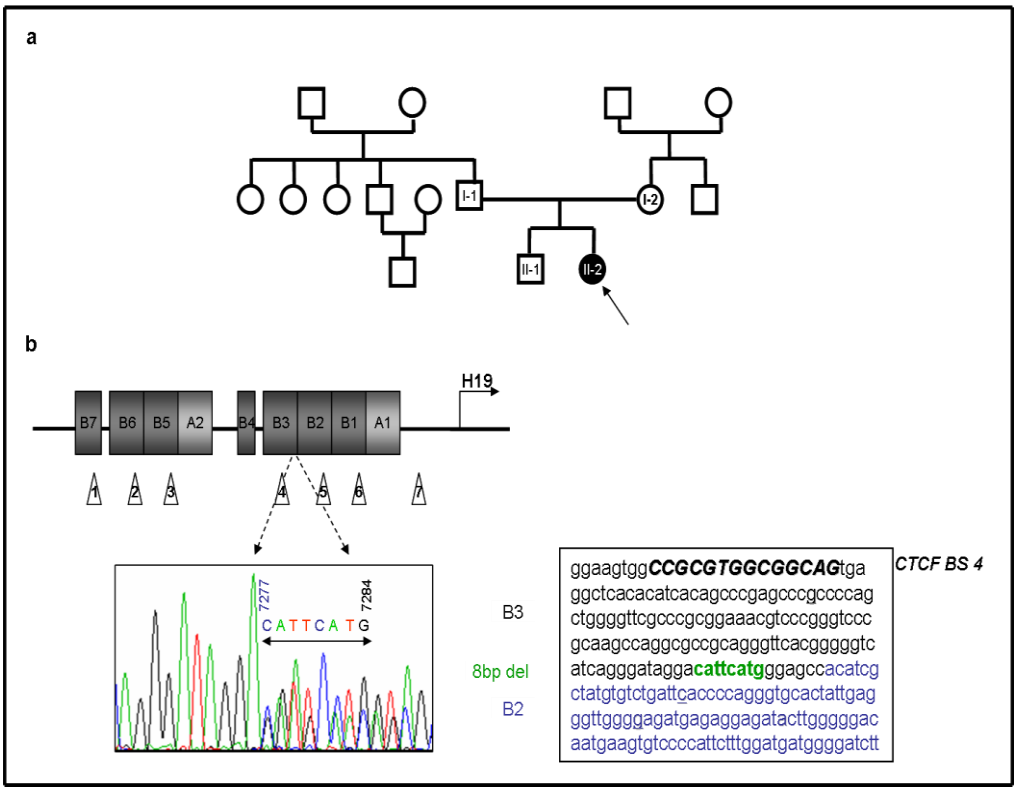


Figure 13: ICR1 deletion in BWS case CF32. **a**) Pedigree of the family. **b**) Identification and localization of the deletion. The 8 bp deletion is located within the B3 domain and breakpoints are indicated by arrows at 7277 and 7284 (according to the reference sequence AF125183). The deletion was outside a CTCF binding site, 116 bases 3' (green letters in the sequences) of the fourth CTCF binding site (capital letters in the sequences) and did not delete any of the seven CTCF binding sites.

ICR1 analysis identified a single nucleotide variation in one SRS patient (PE18), outside a CTCF binding site, (G->T 9048, according to the reference sequence AF 125183). This variation was not found neither in a control population (n=50) nor in all public databases and is therefore unlikely to be a polymorphism (Fig. 14).

Deletions and the single nucleotide variation and their parental inheritance when available are summarized in table VII.

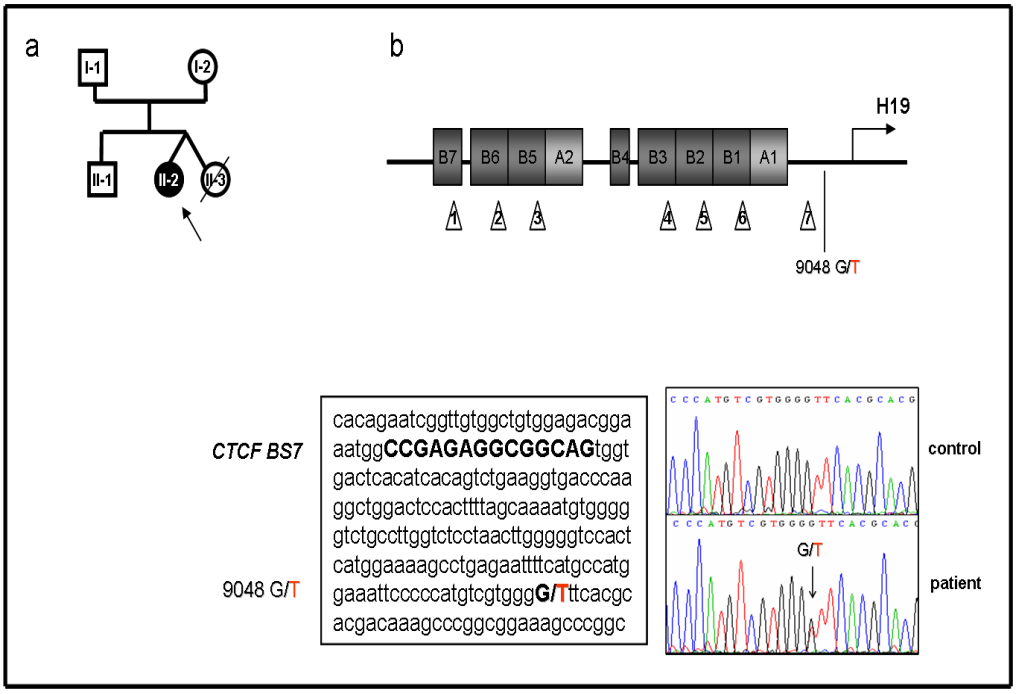


Figure 14: ICR1 analysis in SRS case PE18**. a**) Pedigree of the family. **b**) Identification and localization of the heterogenous variation (G->T 9048, according to the reference sequence AF125183). The variation was outside of a CTCF binding site (CTCF BS), 164 bases 3' from the seventh CTCF binding site (capital letters in the sequences).

Long range PCR and direct sequencing of the seven CTCF binding sites in the familial case (BWS brothers (LB43 and LS44), with gain of methylation at ICR1), ruled out a deletion in this familial case (Fig. 15).

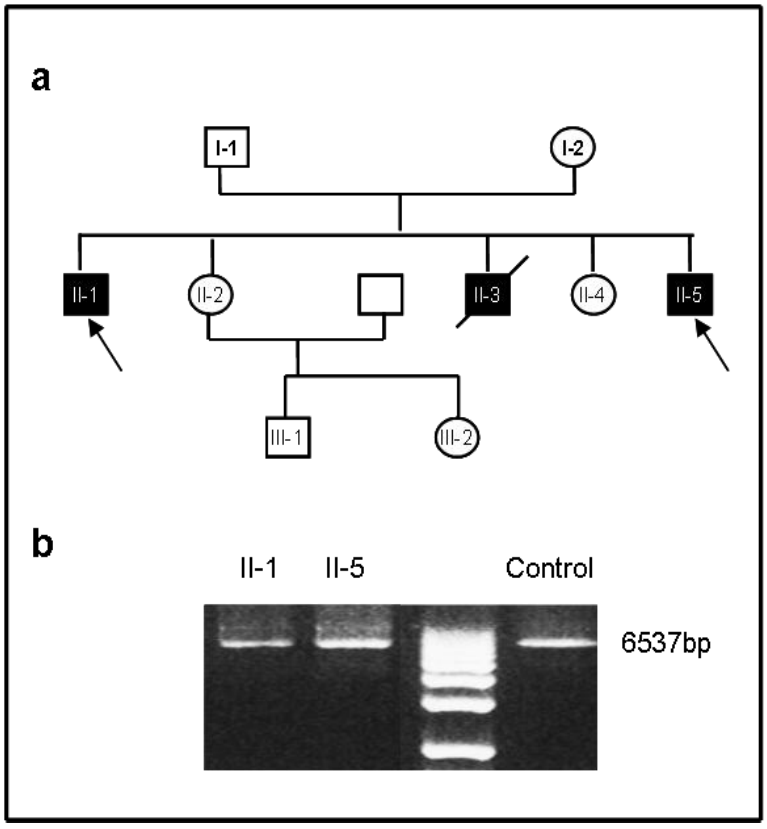


Figure 15: Familial BWS case. **a**) Pedigree of the familial BWS cases. **b**) Long range PCR for the two brothers (II-1 and II-5), and a 0.8% agarose gel ruling out a deletion in this family.

3.1.1.3 Identification of nucleotide variations in the control population

The ICR1 11p15 region is genetically highly variable and displays a number of single-nucleotide polymorphisms (SNPs). Sequencing and analysis of this region in our cohort of patients allowed to identify new genetic changes that were not documented as polymorphisms (position 5269 in a SRS patient (SA21) and 7648 in a BWS patient (TL) according to the reference sequence AF 125183). Therefore, we sequenced a cohort of 50 control samples to evaluate precisely the frequency of these new variations in the normal population (4.0% for the SNP position 5269 and 6.3% for the SNP position 7648). The sequence of control samples also allowed us to document the

frequency of other previously described SNPs, which were not available in public databases. The frequency of the SNPs is summarized in table VIII.

Table VII: Deletions and single nucleotide variation in BWS and SRS patients

Patient	Phenotype	Variation	Position (Acc AF125183)	Localization	Transmission	% Heterozygosity x/n chromosomes
CF32	BWS	del 8 bp	7277-7284	B3/ 3' of CTCFBS4	maternal	0/100
KL41	BWS	Del ~ 1.8 kb	5067/5086-6900/6919	B6>B3	maternal	
PE18	SRS	G->T	9048	3' of CTCFBS7	NA	0/100

NA: not available

Table VIII: Frequency of polymorphisms at the ICR1 region in a control population

SNP number	Variation	Position (Acc. AF125183)	NCBI frequency % heterozygosity	% Heterozygosity in our control population(n=50)
rs57889360	GA	5165	unknown	2.0%
(CERRATO *et al.* 2008)	TC	5198	unknown	8.2%
This study*	TC*	5269	unknown	4.0%
rs4930101	AC	7313	unknown	31.9%
rs2525882	AG	7342	unknown	29.2%
rs2735970	AG	7357	unknown	32.6%
rs2735971	AG	7523	unknown	29.2%
rs2735972	CT	7591	unknown	29.2%
This study*	CA*	7648	unknown	6.3%
(CERRATO *et al.* 2008)	AG	7679	unknown	18.4%

*: new polymorphisms

3.1.2 ICR1 DNA methylation analysis

Since a gain of methylation at ICR1 in patients with ICR1 deletion has been controversial, we decided to analyze the DNA methylation status in our BWS patients carrying an ICR1 deletion (patients CF32 and KL41). In addition, the DNA methylation profile was investigated in the SRS patient with the single nucleotide variation at ICR1 (PE18). In this study, DNA methylation status was analyzed by bisulfite sequencing allowing the analysis of each CpG in the CTCF binding sites and their flanking regions.

Bisulfite sequencing demonstrated a gain of methylation at CTCF binding sites 1, 4, 6 and 7 (methylation indexes from 73% to 95%) in BWS patient KL41 (Fig. 16b). CTCF binding sites 2 and 3 were also investigated in KL 41 and the results indicated that both sites were methylated (98% and 90% respectively) on the wild type paternal allele, because these two sites were abolished on the maternal allele that carries the deletion.

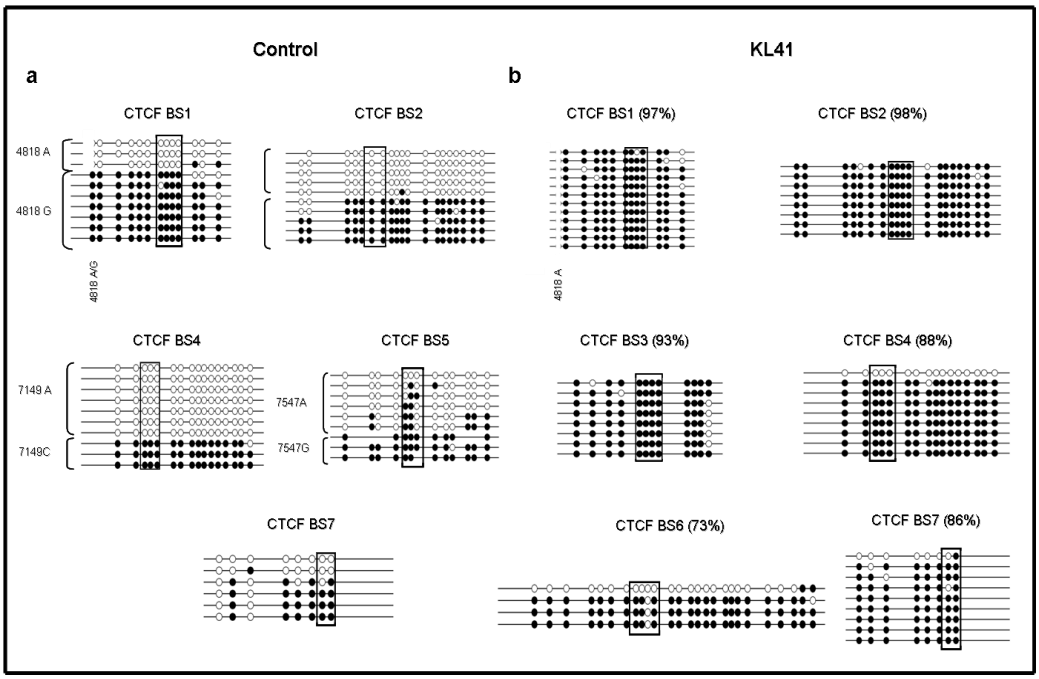


Figure 16: Methylation status of ICR1 in BWS case KL41 and a control individual. **a**) DNA methylation profiles of some CTCF binding sites (CTCF BS) and flanking regions determined by bisulfite sequencing in control sample (in blood cells). Each line corresponds to an individual cloned DNA fragment and each circle to a CpG dinucleotide. Methylated CpGs are indicated by filled circles and unmethylated CpGs by open circles. The CpG included in the CTCF binding sites are framed. Parental alleles were distinguished by known single-nucleotide polymorphisms at CTCF binding sites 1, 4 and 5. **b**) DNA methylation profiles of CTCF binding sites 1-4, 6-7 and flanking regions in KL41 patient (in blood cells). CTCF binding sites 2 and 3 were originated from the paternal allele, because these two sites are abolished on the maternal allele that carried the deletion (All numbers are numbered according to AF125183).

Bisulfite sequencing also showed a complete gain of methylation at CTCF binding sites on the maternal allele for BWS patient CF32. We were able to distinguish the two parental alleles and to show that the maternal deleted allele was methylated (Fig. 17). The extent of methylation was similar (methylation index more than 96% for patient CF32) at the seven CTCF binding sites and the *H19* promoter.

These results indicate that a deletion at ICR1 (even a very small deletion which does not involve a CTCF binding site) results in a gain of methylation at ICR1.

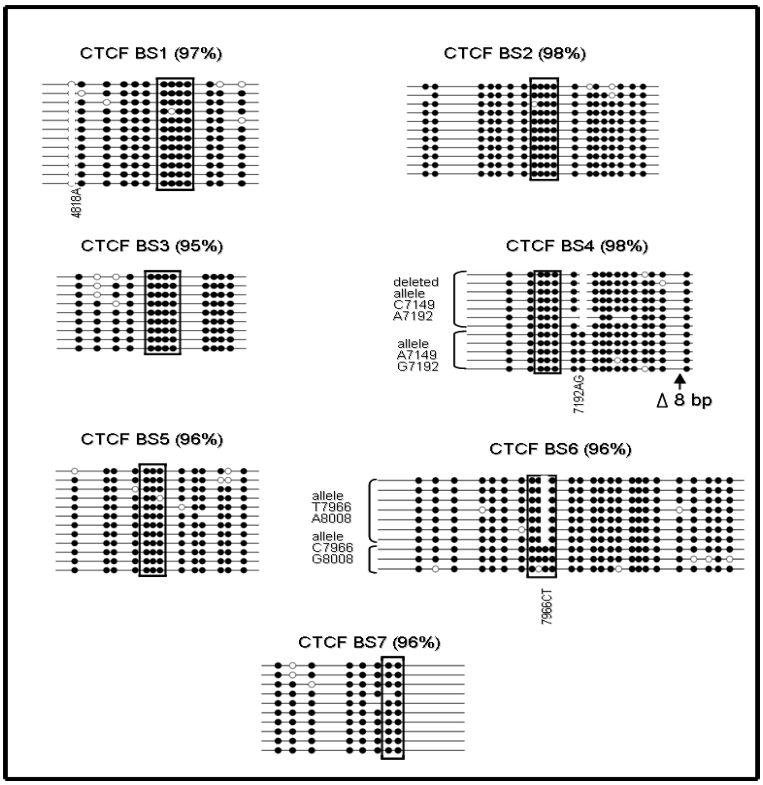


Figure 17: Methylation status of ICR1 in BWS case CF32. DNA methylation pattern in all CTCF binding sites (CTCF BS) 1-7 and flanking regions in patient CF32 (tongue tissue). Maternal (abnormal) and paternal (wild-type) alleles were distinguished by known single-nucleotide polymorphisms (All numbers are numbered according to AF125183).

Conversely, there was a nearly complete loss of methylation on the paternal allele for PE18 patient (SRS patient with the single nucleotide variation) in CTCF binding sites 2, 4 and 6 (Fig. 18) (methylation indexes were between 0% and 4%). A known SNP (7192) in the flanking region of the CTCF binding site 4, showed that both alleles are unmethylated, indicating a loss of methylation on the paternal allele. In the other CTCF binding sites (1, 3 and 5) only one colony in each one of them was methylated and no SNPs in this patient were available to identify if these colonies come from one allele or from different alleles.

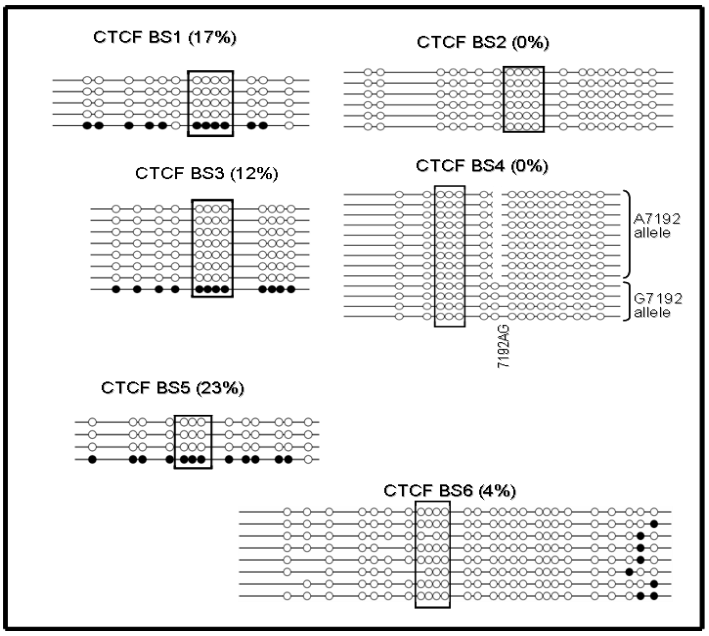


Figure 18: Methylation status of ICR1 in SRS case PE18. DNA methylation profiles of CTCF binding sites 1-6 and flanking regions determined by bisulfite sequencing in blood cells. Paternal allele (with loss of methylation) and Maternal (wild-type) alleles were distinguished by known single-nucleotide polymorphisms (All numbers are numbered according to AF125183).

3.2 Mutation analysis of the *CTCF* gene

CTCF binds the seven CTCF binding sites using its ZFs to prevent the *de novo* methylation at *IGF2/H19*-ICR1. Some *CTCF* mutations have been described in several tumors, including Wilms' tumors (FILIPPOVA *et al.* 2002), which altered CTCF DNA-binding specificity (FILIPPOVA 2008). Therefore, we looked for mutations of the *CTCF* gene, that could affect the structure of ZFs, in all BWS patients with gain of methylation at ICR1 (n=21) (Fig 19). We sequenced all exons and flanking intronic regions of the *CTCF* gene. No change in the coding sequence of the *CTCF* gene was found in any of these patients. Only one of the two polymorphisms, according to the accession number NC_000016, was found in the exon 10 3'UTR (rs6499137, G/T) in 3 patients. These data suggest that *CTCF* gene mutation is not a common cause of gain of methylation at ICR1 in BWS.

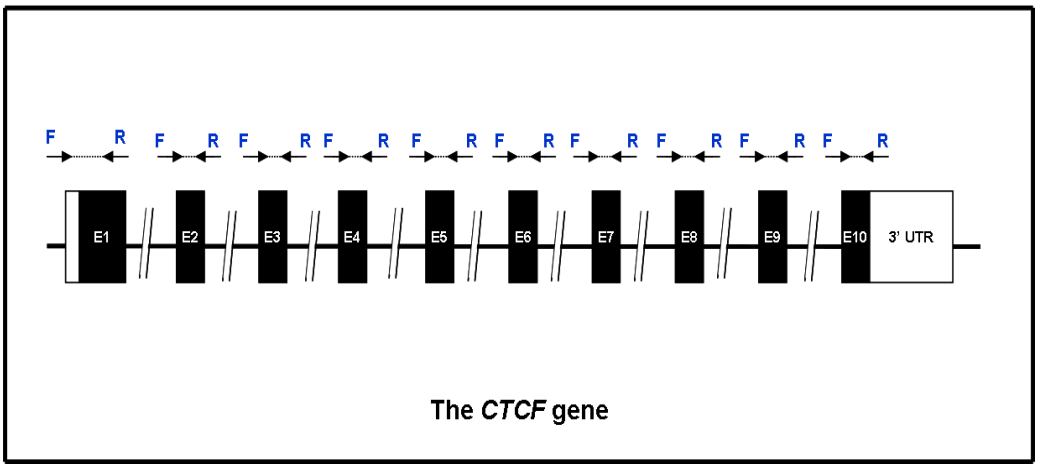


Figure 19: Analysis of the *CTCF* coding sequences. Filled boxes, protein coding sequences (10 exons which were sequenced in all BWS patients); open boxes, untranslated exons; arrows primer sets Forward (F) and reverse (R).

4 DISCUSSION

4 DISCUSSION

As described in the literature, the 11p15 *H19/IGF2*-ICR1 is methylated exclusively on the paternal allele and interacts with the zinc finger protein CTCF on the unmethylated maternal allele at seven CTCF binding sites. The methylation status on the ICR1 mediates the reciprocal imprinting of the two imprinted genes in this locus, *IGF2* and *H19*. Abnormal methylation at ICR1 results in the disturbance of the reciprocal imprinting of these genes and two opposite growth disorders, BWS and SRS. ICR1 is arranged in two blocks which contain highly repetitive sequences. Both the repetition of the sequences and the high content in CpG dinucleotides make this region particularly difficult to analyse. Therefore, the first step of this project was to optimize the PCR conditions. We evaluated different Taq polymerases, different enhancers (betaine, Dimethyl Sulfoxide (DMSO) and commercial enhancers) and different protocols of PCR (touch down) to achieve for each domain the best conditions of amplification.

Abnormal methylation of the 11p15 *IGF2/H19*-ICR1 imprinted domain accounts for approximately 10% of BWS patients (ENKLAAR *et al.* 2006; GASTON *et al.* 2001) and 60% of SRS patients (ROSSIGNOL *et al.* 2008). Little information is available regarding the mechanism of ICR1 DNA methylation defects. Only a few studies have addressed the relationship between ICR1 mutations and DNA methylation defect at *IGF2/H19*-ICR1 locus in humans. Deletions (1.4 to 2.2 kb) removing part of ICR1 have been described in a few BWS familial cases with dominant maternal transmission (CERRATO *et al.* 2008; PRAWITT *et al.* 2005; SPARAGO *et al.* 2004; SPARAGO *et al.* 2007). These deletions do not affect the phenotype when paternally inherited. No smaller deletions or mutations within ICR1 have been described in 12 BWS patients not exhibiting deletions (CERRATO *et al.* 2008) and alter construction of the 11p15.5 haplotypes, the authors suggested that the epimutation in these 12 individuals arose independently of the sequence context (CERRATO *et al.* 2008). Very recently, Scott *et al*.

(SCOTT *et al.* 2008) showed that constitutive ICR1 imprinting abnormalities (including a 5.3 kb deletion and a 2.9 kb insertion) can also occur in apparently sporadic Wilms' tumors. All together, ICR1 deletions account for only a small proportion of BWS cases. Even less is known regarding the mechanism of ICR1 loss of methylation in SRS patients. Engel *et al.* (ENGEL *et al.* 2004) previously showed that mutations of CpG dinucleotides in the mouse *H19* DMD result in a growth retardation phenotype when paternally inherited and loss of methylation of the paternal allele as seen in SRS patients. However, attempts to identify ICR1 mutations in SRS patients, including familial cases, were negative (BARTHOLDI *et al.* 2008; BLIEK *et al.* 2004; BRUCE *et al.* 2008; YAMAZAWA *et al.* 2008).

In this study, we show that 2 of 19 (10%) BWS patients with apparently sporadic BWS, actually display a deletion within ICR1. Long range PCR allowed us to identify a 1.8 kb deletion in one patient (KL41) similar to a deletion described in an unrelated familial case (SPARAGO *et al.* 2004) (SPARAGO *et al.* 2007) (Fig. 12). Additional primers were designed for the ICR1 domain and cloned fragments were bi-directionally sequenced to specifically characterize the 1.8 kb deletion. Another patient displayed a very small deletion involving only 8 bp in the B3 domain, 116 bases 3' of CTCF binding site 4 (Fig. 13). Sandovici *et al.* (SANDOVICI *et al.* 2006) showed that there is an excess of hot-spots of recombination at imprinted loci, particularly at the 11p15 region, compared to the rest of the genome. This higher rate of recombination might explain the incidence and localization of deletions at ICR1. Parental analysis for the two patients with a deletion showed that both deletions were maternally inherited.

The wild-type ICR1 domain is arranged in two repeat blocks and contains seven CTCF target sites. It has remained controversial whether the ICR1 deletion was sufficient to result in gain of DNA methylation (PRAWITT *et al.* 2005; RICCIO 2008). Previously described imprinting centre mutations that cause ICR1 gain of methylation are

deletions that delete 1.4–1.8 kb at the core of ICR1, fusing the two repeat blocks and abolishing one or two CTCF target sites (SPARAGO *et al.* 2004; SPARAGO *et al.* 2007). Intriguingly, another deletion, measuring 2.2 kb, abolishing three CTCF target sites while maintaining one repeat block, did not result in ICR1 gain of methylation (PRAWITT *et al.* 2005). It was suggested at the time that alteration of spacing rather than loss of CTCF target sites might be essential for ICR1 gain of methylation (PRAWITT *et al.* 2005). More recently, a larger deletion of 5.3 kb removing both repeat blocks and six of the seven CTCF binding sites was shown to also result in gain of methylation at the seventh CTCF binding site (SCOTT *et al.* 2008). Surprisingly, an insertion of 2.9 kb adding two CTCF binding sites between the two ICR1 repeat blocks was also shown to result in gain of methylation despite no loss of CTCF target sites or disturbance of the architecture within each block (SCOTT *et al.* 2008).

In order to investigate the consequences of the deletions in our patients with the deletions (in particular the small 8 bp deletion) on the DNA methylation status at ICR1, bisulfite sequencing for CTCF binding sites and their flanking regions has been performed. Bisulfite sequencing of DNA allowed us to precisely analyze each CpG dinucleotides at CTCF binding sites and their flanking regions.

In BWS KL41 case, the deletion (1.8 kb) that abolished CTCF binding sites 2 and 3, resulted in a gain of methylation at CTCF binding sites 1, 4 and 6-7 (methylation indexes were more than 70%) (Fig. 16b). In the 8 bp deletion case, the gain of DNA methylation was nearly complete in tongue tissue and the extent of gain of methylation was homogenous with methylation indexes in the same range at the various CTCF binding sites (Fig. 17). The deletion did not remove any of CTCF binding sites. Although, this deletion was very small and did not alter the structure of the ICR1 (in terms of number of repeated sequences and the spacing of CTCF binding sites), it resulted in a gain of

methylation at ICR1. Thus, the small 8 bp deletion described here does not really support a CTCF binding sites spacing hypothesis.

In the previous reports, ICR1 gain of methylation associated with insertion or deletion had a mosaic pattern and affected only a proportion of mutated cells (CERRATO *et al.* 2008; SCOTT *et al.* 2008; SPARAGO *et al.* 2007). In this study, our data are a little different with a nearly complete gain of methylation at all CTCF binding sites in the two patients with a deletion.

Previous reports have also shown that the level of methylation in patients with ICR1 deletions is different between the proximal (B7-A2) and distal (B4-A1) CTCF binding sites (CERRATO *et al.* 2008; SPARAGO *et al.* 2007). In our study, we showed that the level of methylation is similar at all CTCF binding sites.

No abnormality at ICR1 or the *CTCF* gene has been found in the familial cases (Fig. 15). This family is very special in the way that three brothers displayed an overgrowth phenotype when two sisters were unaffected. This familial pedigree was evocative of an X linked disorder. As there is an X linked overgrowth disorder that shares some phenotypes with BWS, this family was initially investigated for the X linked Simpson-Golabi-Behmel syndrome (PILIA *et al.* 1996). The sequencing of the *Glypican 3* gene was normal. Instead, analysis of the 11p15 region showed that the affected brothers displayed a gain of methylation at ICR1. The father who was born with macrosomia and displays a supernumerary nipple had a normal methylation profile. At this stage the mechanism of the methylation defect in this family remains unknown.

In order to identify the cause of the epigenetic defect in the SRS patients with loss of methylation at ICR1, a long range PCR of ICR1 and direct sequencing of the CTCF binding sites and their flanking region have been conducted. No deletions have been found in SRS

cases. However, a single nucleotide variation was found in a non familial SRS case (Fig. 14). This heterozygous variation was outside of CTCF binding sequences, 3' of CTCF binding site 7. This single nucleotide variation is not a known SNP and was not found in 50 control samples. Bisulfite sequencing showed a loss of methylation at CTCF binding sites. A SNP within the 4th CTCF binding site allowed us to confirm that there was a loss of methylation on the paternal allele (Fig. 18). At this stage, we can not be sure if this loss of methylation results from the mutation or if the loss of methylation results from another unknown mechanism like in most SRS patients. Unfortunately, no cells or tissues are available for this patient to test, by chromatin immunoprecipitation or chromosome conformation capture, the relevance of the mutation.

The sequence of the *CTCF* gene was normal for all BWS patients including the familial cases. Mutations of the *CTCF* gene have been previously described in several tumors including two Wilms' tumors (FILIPPOVA *et al.* 2002). Wilms' tumor is very common in BWS patients, and more particularly in BWS with a gain of methylation at ICR1 (ENKLAAR *et al.* 2006). However, other studies failed to find any CTCF mutation in Wilms' tumors (CUI *et al.* 2001; YEH *et al.* 2002). Regarding the role of CTCF in the maintenance of the hypomethylated state of the maternal allele, we hypothesized that deleterious mutation of *CTCF* might account for some BWS with ICR1 gain of methylation. No mutation was found and our data are consistent with a previous publication (CERRATO *et al.* 2008).

The ICR1 11p15 region is genetically highly variable and displays a lot of SNPs. When we sequenced this region in our cohort of patients, we described new genetic changes (2 SNPs) that were not found in more than one patient and were not documented as polymorphisms. We evaluated precisely the frequency of the two new SNPs in the normal population (positions 5269 and 7648 according to the reference

sequence AF 125183) and we also documented the frequency of other previously described SNPs (table VIII).

In conclusion, we described two new BWS cases and one SRS case with ICR1 genetic defect. We show that a very small 8 bp deletion which does not delete any of the CTCF binding sites results in a gain at methylation of all CTCF binding sites. Although accounting for a very small proportion of BWS patients, these deletions should be recognized. BWS patients with ICR1 gain of methylation have a high risk of tumors (mainly Wilms' tumors) and their molecular diagnosis and the recognition of a deletion is particularly important for genetic counseling and tumor surveillance.

At this stage, the mechanism of the methylation defect at ICR1 remains largely unknown in both BWS and SRS patients. By opposition to BWS patients with ICR2 loss of methylation, gain of DNA methylation is localized to the *H19/IGF2* ICR1 domain and does not affect other paternally- or maternally methylated loci (BLIEK *et al.* 2008). Recent studies have demonstrated that other cofactors such as cohesin (WENDT *et al.* 2008), MBD3 (REESE *et al.* 2007), the EZH2 histone methyltransferase (LI *et al.* 2008) or *91H*, a 120 kb *H19* antisense RNA (BERTEAUX *et al.* 2008) participate to the regulation of imprinting and chromatin looping at the *H19/IGF2* ICR1 domain but their putative role in imprinting disorders have not been addressed yet. Alternatively, other recent studies suggest that particular genotypes at 11p15 might affect the susceptibility to epigenetic signals (HEIJMANS *et al.* 2007) and extensive future studies should also address this question.

BIBLIOGRAPHY

BAKER, J., J. P. LIU, E. J. ROBERTSON and A. EFSTRATIADIS, 1993 Role of insulin-like growth factors in embryonic and postnatal growth. Cell **75:** 73-82.

BARLOW, D. P., R. STOGER, B. G. HERRMANN, K. SAITO and N. SCHWEIFER, 1991 The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature **349:** 84-87.

BARTHOLDI, D., M. KRAJEWSKA-WALASEK, K. OUNAP, H. GASPAR, K. H. CHRZANOWSKA *et al.*, 2008 Epigenetic mutations of the imprinted IGF2-H19 domain in Silver-Russell Syndrome (SRS): Results from a large cohort of patients with SRS and SRS-like phenotypes. J Med Genet.

BARTOLOMEI, M. S., S. ZEMEL and S. M. TILGHMAN, 1991 Parental imprinting of the mouse H19 gene. Nature **351:** 153-155.

BELL, A. C., A. G. WEST and G. FELSENFELD, 1999 The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell **98:** 387-396.

BERGER, S. L., 2007 The complex language of chromatin regulation during transcription. Nature **447:** 407-412.

BERTEAUX, N., N. APTEL, G. CATHALA, C. GENTON, J. COLL *et al.*, 2008 A novel H19 antisense RNA overexpressed in breast cancer contributes to paternal IGF2 expression. Mol Cell Biol **28:** 6731-6745.

BESTOR, T. H., 2000 The DNA methyltransferases of mammals. Hum Mol Genet **9:** 2395-2402.

BLIEK, J., C. GICQUEL, S. MAAS, V. GASTON, Y. LE BOUC *et al.*, 2004 Epigenotyping as a tool for the prediction of tumor risk and tumor type in patients with Beckwith-Wiedemann syndrome (BWS). J Pediatr **145:** 796-799.

BLIEK, J., P. TERHAL, M. J. VAN DEN BOGAARD, S. MAAS, B. HAMEL *et al.*, 2006 Hypomethylation of the H19 gene causes not only Silver-Russell syndrome (SRS) but also isolated asymmetry or an SRS-like phenotype. Am J Hum Genet **78:** 604-614.

BLIEK, J., G. VERDE, J. CALLAWAY, S. M. MAAS, A. DE CRESCENZO *et al.*, 2008 Hypomethylation at multiple maternally methylated imprinted

regions including PLAGL1 and GNAS loci in Beckwith-Wiedemann syndrome. Eur J Hum Genet.

BOURC'HIS, D., G. L. XU, C. S. LIN, B. BOLLMAN and T. H. BESTOR, 2001 Dnmt3L and the establishment of maternal genomic imprints. Science **294:** 2536-2539.

BRANNAN, C. I., E. C. DEES, R. S. INGRAM and S. M. TILGHMAN, 1990 The product of the H19 gene may function as an RNA. Mol Cell Biol **10:** 28-36.

BRUCE, S., K. HANNULA-JOUPPI, C. M. LINDGREN, M. LIPSANEN-NYMAN and J. KERE, 2008 Restriction site-specific methylation studies of imprinted genes with quantitative real-time PCR. Clin Chem **54:** 491-499.

CAI, X., and B. R. CULLEN, 2007 The imprinted H19 noncoding RNA is a primary microRNA precursor. RNA **13:** 313-316.

CERRATO, F., A. SPARAGO, G. VERDE, A. DE CRESCENZO, V. CITRO *et al.*, 2008 Different Mechanisms Cause Imprinting defects at the IGF2/H19 locus in Beckwith-Wiedemann Syndrome and Wilms' Tumour. Hum Mol Genet.

CHARALAMBOUS, M., T. R. MENHENIOTT, W. R. BENNETT, S. M. KELLY, G. DELL *et al.*, 2004 An enhancer element at the Igf2/H19 locus drives gene expression in both imprinted and non-imprinted tissues. Dev Biol **271:** 488-497.

CONSTANCIA, M., M. HEMBERGER, J. HUGHES, W. DEAN, A. FERGUSON-SMITH *et al.*, 2002 Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature **417:** 945-948.

COOPER, W. N., A. LUHARIA, G. A. EVANS, H. RAZA, A. C. HAIRE *et al.*, 2005 Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. Eur J Hum Genet **13:** 1025-1032.

CUI, H., E. L. NIEMITZ, J. D. RAVENEL, P. ONYANGO, S. A. BRANDENBURG *et al.*, 2001 Loss of imprinting of insulin-like growth factor-II in Wilms' tumor commonly involves altered methylation but not mutations of CTCF or its binding site. Cancer Res **61:** 4947-4950.

DECHIARA, T. M., E. J. ROBERTSON and A. EFSTRATIADIS, 1991 Parental imprinting of the mouse insulin-like growth factor II gene. Cell **64:** 849-859.

DELAVAL, K., and R. FEIL, 2004 Epigenetic regulation of mammalian genomic imprinting. Curr Opin Genet Dev **14:** 188-195.

DELAVAL, K., J. GOVIN, F. CERQUEIRA, S. ROUSSEAUX, S. KHOCHBIN *et al.*, 2007 Differential histone modifications mark mouse imprinting control regions during spermatogenesis. EMBO J **26:** 720-729.

DELAVAL, K., A. WAGSCHAL and R. FEIL, 2006 Epigenetic deregulation of imprinting in congenital diseases of aberrant growth. Bioessays **28:** 453-459.

DHASARATHY, A., and P. A. WADE, 2008 The MBD protein family-reading an epigenetic mark? Mutat Res **647:** 39-43.

DREWELL, R. A., C. J. GODDARD, J. O. THOMAS and M. A. SURANI, 2002 Methylation-dependent silencing at the H19 imprinting control region by MeCP2. Nucleic Acids Res **30:** 1139-1144.

ENGEL, N., J. L. THORVALDSEN and M. S. BARTOLOMEI, 2006 CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. Hum Mol Genet **15:** 2945-2954.

ENGEL, N., A. G. WEST, G. FELSENFELD and M. S. BARTOLOMEI, 2004 Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. Nat Genet **36:** 883-888.

ENKLAAR, T., B. U. ZABEL and D. PRAWITT, 2006 Beckwith-Wiedemann syndrome: multiple molecular mechanisms. Expert Rev Mol Med **8:** 1-19.

FEDORIW, A. M., P. STEIN, P. SVOBODA, R. M. SCHULTZ and M. S. BARTOLOMEI, 2004 Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science **303:** 238-240.

FILIPPOVA, G. N., 2008 Genetics and epigenetics of the multifunctional protein CTCF. Curr Top Dev Biol **80:** 337-360.

FILIPPOVA, G. N., S. FAGERLIE, E. M. KLENOVA, C. MYERS, Y. DEHNER *et al.*, 1996 An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol Cell Biol **16:** 2802-2813.

FILIPPOVA, G. N., C. F. QI, J. E. ULMER, J. M. MOORE, M. D. WARD *et al.*, 2002 Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. Cancer Res **62:** 48-52.

FOWDEN, A. L., C. SIBLEY, W. REIK and M. CONSTANCIA, 2006 Imprinted genes, placental development and fetal growth. Horm Res **65 Suppl 3:** 50-58.

FREVEL, M. A., S. J. SOWERBY, G. B. PETERSEN and A. E. REEVE, 1999 Methylation sequencing analysis refines the region of H19 epimutation in Wilms tumor. J Biol Chem **274:** 29331-29340.

GABORY, A., M. A. RIPOCHE, T. YOSHIMIZU and L. DANDOLO, 2006 The H19 gene: regulation and function of a non-coding RNA. Cytogenet Genome Res **113:** 188-193.

GASTON, V., Y. LE BOUC, V. SOUPRE, L. BURGLEN, J. DONADIEU *et al.*, 2001 Analysis of the methylation status of the KCNQ1OT and H19 genes in leukocyte DNA for the diagnosis and prognosis of Beckwith-Wiedemann syndrome. Eur J Hum Genet **9:** 409-418.

GICQUEL, C., and Y. LE BOUC, 2006 Hormonal regulation of fetal growth. Horm Res **65 Suppl 3:** 28-33.

GICQUEL, C., Y. LE BOUC and A. EL-OSTA, 2008 Epigenetic regulation and fetal programming. Best Practice & Research Clinical Endocrinology & Metabolism **22, No. 1:** 1-16.

GICQUEL, C., S. ROSSIGNOL, S. CABROL, M. HOUANG, V. STEUNOU *et al.*, 2005 Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. Nat Genet **37:** 1003-1007.

HAJKOVA, P., S. ERHARDT, N. LANE, T. HAAF, O. EL-MAARRI *et al.*, 2002 Epigenetic reprogramming in mouse primordial germ cells. Mech Dev **117:** 15-23.

HAN, L., D. H. LEE and P. E. SZABO, 2008 CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. Mol Cell Biol **28:** 1124-1135.

HATA, K., M. OKANO, H. LEI and E. LI, 2002 Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development **129:** 1983-1993.

HEIJMANS, B. T., D. KREMER, E. W. TOBI, D. I. BOOMSMA and P. E. SLAGBOOM, 2007 Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum Mol Genet **16:** 547-554.

HOWELL, C. Y., T. H. BESTOR, F. DING, K. E. LATHAM, C. MERTINEIT *et al.*, 2001 Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell **104:** 829-838.

IDERAABDULLAH, F. Y., S. VIGNEAU and M. S. BARTOLOMEI, 2008 Genomic imprinting mechanisms in mammals. Mutat Res **647:** 77-85.

KERJEAN, A., J. M. DUPONT, C. VASSEUR, D. LE TESSIER, L. CUISSET *et al.*, 2000 Establishment of the paternal methylation imprint of the human H19 and MEST/PEG1 genes during spermatogenesis. Hum Mol Genet **9:** 2183-2187.

KOUZARIDES, T., 2007 Chromatin modifications and their function. Cell **128:** 693-705.

KRUEGER, C., and C. S. OSBORNE, 2006 Raising the curtains on interchromosomal interactions. Trends Genet **22:** 637-639.

LEWIS, A., K. MITSUYA, D. UMLAUF, P. SMITH, W. DEAN *et al.*, 2004 Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. Nat Genet **36:** 1291-1295.

LEWIS, A., and W. REIK, 2006 How imprinting centres work. Cytogenet Genome Res **113:** 81-89.

LI, B., M. CAREY and J. L. WORKMAN, 2007 The role of chromatin during transcription. Cell **128:** 707-719.

LI, E., 2002 Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet **3:** 662-673.

LI, E., T. H. BESTOR and R. JAENISCH, 1992 Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **69:** 915-926.

LI, T., J. F. HU, X. QIU, J. LING, H. CHEN *et al.*, 2008 CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. Mol Cell Biol **28:** 6473-6482.

LING, J. Q., T. LI, J. F. HU, T. H. VU, H. L. CHEN *et al.*, 2006 CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science **312:** 269-272.

LOPES, S., A. LEWIS, P. HAJKOVA, W. DEAN, J. OSWALD *et al.*, 2003 Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. Hum Mol Genet **12:** 295-305.

MAGER, J., N. D. MONTGOMERY, F. P. DE VILLENA and T. MAGNUSON, 2003 Genome imprinting regulated by the mouse Polycomb group protein Eed. Nat Genet **33:** 502-507.

MCGRATH, J., and D. SOLTER, 1984 Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell **37:** 179-183.

MILLER, S. A., D. D. DYKES and H. F. POLESKY, 1988 A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res **16:** 1215.

MONK, D., R. SANCHES, P. ARNAUD, S. APOSTOLIDOU, F. A. HILLS *et al.*, 2006 Imprinting of IGF2 P0 transcript and novel alternatively spliced INS-IGF2 isoforms show differences between mouse and human. Hum Mol Genet **15:** 1259-1269.

MOON, H., G. FILIPPOVA, D. LOUKINOV, E. PUGACHEVA, Q. CHEN *et al.*, 2005 CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep **6:** 165-170.

MORGAN, H. D., F. SANTOS, K. GREEN, W. DEAN and W. REIK, 2005 Epigenetic reprogramming in mammals. Hum Mol Genet **14 Spec No 1:** R47-58.

MURRELL, A., S. HEESON and W. REIK, 2004 Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat Genet **36:** 889-893.

MURRELL, A., Y. ITO, G. VERDE, J. HUDDLESTON, K. WOODFINE *et al.*, 2008 Distinct methylation changes at the IGF2-H19 locus in congenital growth disorders and cancer. PLoS ONE **3:** e1849.

OHLSSON, R., F. HEDBORG, L. HOLMGREN, C. WALSH and T. J. EKSTROM, 1994 Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. Development **120:** 361-368.

OHLSSON, R., A. NYSTROM, S. PFEIFER-OHLSSON, V. TOHONEN, F. HEDBORG *et al.*, 1993 IGF2 is parentally imprinted during human embryogenesis and in the Beckwith-Wiedemann syndrome. Nat Genet **4:** 94-97.

OHLSSON, R., R. RENKAWITZ and V. LOBANENKOV, 2001 CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet **17:** 520-527.

OKANO, M., D. W. BELL, D. A. HABER and E. LI, 1999 DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell **99:** 247-257.

OZANNE, S. E., and M. CONSTANCIA, 2007 Mechanisms of disease: the developmental origins of disease and the role of the epigenotype. Nat Clin Pract Endocrinol Metab **3:** 539-546.

PANT, V., P. MARIANO, C. KANDURI, A. MATTSSON, V. LOBANENKOV *et al.*, 2003 The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the H19 imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains. Genes Dev **17:** 586-590.

PILIA, G., R. M. HUGHES-BENZIE, A. MACKENZIE, P. BAYBAYAN, E. Y. CHEN *et al.*, 1996 Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. Nat Genet **12:** 241-247.

PRAWITT, D., T. ENKLAAR, B. GARTNER-RUPPRECHT, C. SPANGENBERG, M. OSWALD *et al.*, 2005 Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. Proc Natl Acad Sci U S A **102:** 4085-4090.

REESE, K. J., S. LIN, R. I. VERONA, R. M. SCHULTZ and M. S. BARTOLOMEI, 2007 Maintenance of paternal methylation and repression of the imprinted H19 gene requires MBD3. PLoS Genet **3:** e137.

REIK, W., and J. WALTER, 2001 Genomic imprinting: parental influence on the genome. Nat Rev Genet **2:** 21-32.

RICCIO, A., 2008 Wilms tumor and constitutional epigenetic defects. Nat Genet **40:** 1272-1273.

ROSSIGNOL, S., I. NETCHINE, Y. LE BOUC and C. GICQUEL, 2008 Epigenetics in Silver-Russell syndrome. Best Pract Res Clin Endocrinol Metab **22:** 403-414.

ROYO, H., and J. CAVAILLE, 2008 Non-coding RNAs in imprinted gene clusters. Biol Cell **100:** 149-166.

SAKATANI, T., A. KANEDA, C. A. IACOBUZIO-DONAHUE, M. G. CARTER, S. DE BOOM WITZEL *et al.*, 2005 Loss of imprinting of Igf2 alters intestinal maturation and tumorigenesis in mice. Science **307:** 1976-1978.

SANDOVICI, I., S. KASSOVSKA-BRATINOVA, J. E. VAUGHAN, R. STEWART, M. LEPPERT *et al.*, 2006 Human imprinted chromosomal regions are historical hot-spots of recombination. PLoS Genet **2:** e101.

SARRAF, S. A., and I. STANCHEVA, 2004 Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol Cell **15:** 595-605.

SCOTT, R. H., J. DOUGLAS, L. BASKCOMB, N. HUXTER, K. BARKER *et al.*, 2008 Constitutional 11p15 abnormalities, including heritable imprinting center mutations, cause nonsyndromic Wilms tumor. Nat Genet **40:** 1329-1334.

SHA, K., 2008 A mechanistic view of genomic imprinting. Annu Rev Genomics Hum Genet **9:** 197-216.

SIMON, I., T. TENZEN, B. E. REUBINOFF, D. HILLMAN, J. R. MCCARREY *et al.*, 1999 Asynchronous replication of imprinted genes is established in the gametes and maintained during development. Nature **401:** 929-932.

SOLTER, D., 2006 Imprinting today: end of the beginning or beginning of the end? Cytogenet Genome Res **113:** 12-16.

SPARAGO, A., F. CERRATO, M. VERNUCCI, G. B. FERRERO, M. C. SILENGO *et al.*, 2004 Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. Nat Genet **36:** 958-960.

SPARAGO, A., S. RUSSO, F. CERRATO, S. FERRAIUOLO, P. CASTORINA *et al.*, 2007 Mechanisms causing imprinting defects in familial Beckwith-

Wiedemann syndrome with Wilms' tumour. Hum Mol Genet **16:** 254-264.

STRAHL, B. D., and C. D. ALLIS, 2000 The language of covalent histone modifications. Nature **403:** 41-45.

SURANI, M. A., S. C. BARTON and M. L. NORRIS, 1984 Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature **308:** 548-550.

TADA, T., M. TADA, K. HILTON, S. C. BARTON, T. SADO *et al.*, 1998 Epigenotype switching of imprintable loci in embryonic germ cells. Dev Genes Evol **207:** 551-561.

TAKAI, D., F. A. GONZALES, Y. C. TSAI, M. J. THAYER and P. A. JONES, 2001 Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. Hum Mol Genet **10:** 2619-2626.

TRASLER, J. M., 2006 Gamete imprinting: setting epigenetic patterns for the next generation. Reprod Fertil Dev **18:** 63-69.

TRASLER, J. M., D. G. TRASLER, T. H. BESTOR, E. LI and F. GHIBU, 1996 DNA methyltransferase in normal and Dnmtn/Dnmtn mouse embryos. Dev Dyn **206:** 239-247.

TURKER, M. S., 1999 The establishment and maintenance of DNA methylation patterns in mouse somatic cells. Semin Cancer Biol **9:** 329-337.

ULANER, G. A., Y. YANG, J. F. HU, T. LI, T. H. VU *et al.*, 2003 CTCF binding at the insulin-like growth factor-II (IGF2)/H19 imprinting control region is insufficient to regulate IGF2/H19 expression in human tissues. Endocrinology **144:** 4420-4426.

UMLAUF, D., Y. GOTO, R. CAO, F. CERQUEIRA, A. WAGSCHAL *et al.*, 2004 Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. Nat Genet **36:** 1296-1300.

WEIDMAN, R. L. J. A. J. R., 2007 Imprinted and More Equal. American Scientist **95:** 143 -149

WENDT, K. S., K. YOSHIDA, T. ITOH, M. BANDO, B. KOCH *et al.*, 2008 Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature **451:** 796-801.

WILKINSON, L. S., W. DAVIES and A. R. ISLES, 2007 Genomic imprinting effects on brain development and function. Nat Rev Neurosci **8:** 832-843.

WORKMAN, J. L., and R. E. KINGSTON, 1998 Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annu Rev Biochem **67:** 545-579.

YAMAZAWA, K., M. KAGAMI, M. FUKAMI, K. MATSUBARA and T. OGATA, 2008 Monozygotic female twins discordant for Silver-Russell syndrome and hypomethylation of the H19-DMR. J Hum Genet **53:** 950-955.

YEH, A., M. WEI, S. B. GOLUB, D. J. YAMASHIRO, V. V. MURTY *et al.*, 2002 Chromosome arm 16q in Wilms tumors: unbalanced chromosomal translocations, loss of heterozygosity, and assessment of the CTCF gene. Genes Chromosomes Cancer **35:** 156-163.

YU, W., V. GINJALA, V. PANT, I. CHERNUKHIN, J. WHITEHEAD *et al.*, 2004 Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. Nat Genet **36:** 1105-1110.

ZHANG, Y., and B. TYCKO, 1992 Monoallelic expression of the human H19 gene. Nat Genet **1:** 40-44.